

# Lanczos-Lovelock models of gravity

T. Padmanabhan\*,

IUCAA, Post Bag 4, Ganeshkhind, Pune - 411 007, India

D. Kothawala<sup>†</sup>

Department of Physics, IIT Madras, Chennai - 600 036, India.

February 13, 2013

## Abstract

Lanczos-Lovelock models of gravity represent a natural and elegant generalization of Einstein's theory of gravity to higher dimensions. They are characterized by the fact that the field equations only contain up to second derivatives of the metric even though the action functional can be a quadratic or higher degree polynomial in the curvature tensor. Because these models share several key properties of Einstein's theory they serve as a useful set of candidate models for testing the emergent paradigm for gravity. This review highlights several geometrical and thermodynamical aspects of Lanczos-Lovelock models which have attracted recent attention.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Lanczos-Lovelock models: Structural Aspects</b>	<b>3</b>
2.1	Lagrangians from curvature and metric: general results . . . . .	4
2.2	Lanczos-Lovelock action functional . . . . .	6
2.3	Explicit form of the equations of motion . . . . .	10
2.4	Diffeomorphism invariance and the Noether current . . . . .	13
2.5	Action functional in the critical dimension . . . . .	14
2.6	Boundary terms for the action principle . . . . .	17
2.7	Holographic structure of Lanczos-Lovelock action . . . . .	20
2.7.1	Higher derivative actions yielding second order equations of motion . .	20
2.7.2	Holography of gravitational action functionals . . . . .	22
2.8	Black hole and cosmological solutions of Lanczos-Lovelock models . . . . .	24

---

\*paddy@iucaa.ernet.in

<sup>†</sup>dawood@physics.iitm.ac.in

<b>3</b>	<b>Lanczos-Lovelock models: Horizon thermodynamics</b>	<b>25</b>
3.1	Horizon entropy in Lanczos-Lovelock theory . . . . .	25
3.2	Entropy of self-gravitating systems and Wald entropy . . . . .	30
3.3	Membrane paradigm in Lanczos-Lovelock models . . . . .	37
3.3.1	Membrane paradigm: summary of the results . . . . .	40
3.3.2	Membrane paradigm: details of the analysis . . . . .	41
3.4	Second law for black hole entropy in Lanczos-Lovelock models . . . . .	45
3.5	Entropy from horizon Virasoro algebra . . . . .	49
<b>4</b>	<b>Lanczos-Lovelock models and the emergent gravity paradigm</b>	<b>55</b>
4.1	Emergent aspects of the Lanczos-Lovelock equations of motion . . . . .	57
4.1.1	Near-horizon symmetries of equation of motion . . . . .	57
4.1.2	Thermodynamic structure of Lanczos-Lovelock field equations . . . .	62
4.1.3	Field equations from the Clausius relation . . . . .	65
4.2	Equipartition law for microscopic degrees of freedom . . . . .	69
4.2.1	Equipartition law in Einstein theory . . . . .	69
4.2.2	Equipartition law in Lanczos-Lovelock models . . . . .	70
4.2.3	Holography and quantization of gravitational entropy . . . . .	71
4.3	Emergent aspects of the Lanczos-Lovelock action . . . . .	74
4.3.1	Lanczos-Lovelock action as free energy of spacetime . . . . .	75
<b>5</b>	<b>Entropy density of spacetime and Lanczos-Lovelock models</b>	<b>77</b>
<b>6</b>	<b>Conclusions</b>	<b>80</b>
<b>A</b>	<b>Algebraic aspects of certain tensors in higher curvature theories of gravity</b>	<b>81</b>
<b>B</b>	<b>Construction of the Gauss-Bonnet lagrangian</b>	<b>85</b>
<b>C</b>	<b>Proof of identities in Eqs. (146) - (148)</b>	<b>86</b>

## 1 Introduction

The principle of equivalence, along with the principle of general covariance, suggests that gravity has a geometrical interpretation in terms of the spacetime structure. These principles are adequate to determine the *kinematics* of gravity (viz., how curvature of spacetime influences the dynamics of spacetime) in an elegant manner. Much of the beauty attributed to general theory of relativity arises from this geometrical interpretation and the kinematics alone.

To complete the picture, one also needs to determine the *dynamics*, viz., how matter determines the spacetime structure. Based on the structure of other physical theories, it appears reasonable to expect that this should arise through a second order differential equation relating the metric tensor to suitably defined source built from matter variables. Unfortunately, we do not have any equally elegant principle to determine these field equations of gravity.

It is therefore reasonable to study the most general system of field equations which are second order in the independent variables, which are the components of the metric tensor. Under very reasonable assumptions, we are then led to a class of models first discussed by Lanczos [1] and Lovelock [2]. These field equations follow from an action functional which is a *polynomial* in the curvature tensor and reduces — uniquely — to Einstein’s general relativity when  $D = 4$ . In higher dimensions ( $D > 4$ ) the Lanczos-Lovelock models differ from Einstein’s theory and have a richer structure.

The purpose of this review is to describe several features of the Lanczos-Lovelock models. Given the fact that they are more general than Einstein’s theory and have been the subject of intensive study in the recent years, it is not possible to cover all the varied aspects of these models. In order to retain focus in the review, we will concentrate on the aspects of Lanczos-Lovelock models closely related to the emergent paradigm for gravity [3, 4, 5]. It turns out that several features related to the thermodynamics of horizons carry over from Einstein’s general relativity to Lanczos-Lovelock models. Since the emergent paradigm of gravity uses the thermodynamic features of horizons as a key input, this allows a natural extension of this paradigm to all Lanczos-Lovelock models. The discovery that several features of the emergent paradigm transcends Einstein’s theory of gravity strongly suggests that the relationship between gravitational dynamics and horizon thermodynamics is more general than was first recognized. In fact, this has been a strong motivation to study Lanczos-Lovelock models of gravity in the recent years. Such studies have also unravelled several peculiar geometrical features of Lanczos-Lovelock models which deserve attention on their own. We will describe these features as well, as we go along.

Unless otherwise specified, we work with a  $D$  dimensional spacetime with signature  $(-, +, +, \dots)$ . Latin alphabets  $i, j, \dots$  run over spacetime indices, whereas Greek symbols  $\mu, \nu, \dots$  run over indices representing a  $(D - 1)$  dimensional hypersurface. This would be usually a spacelike hypersurface, except in Sections 3.2, 3.3 where it is timelike. Finally, we use upper case Latin letters  $A, B, \dots$  for indices representing a  $(D - 2)$  dimensional spacelike hypersurface. The convention for Riemann tensor we follow is:  $R^a_{bcd} = 2\partial_{[c}\Gamma^a_{d]b} + 2\Gamma^a_{m[c}\Gamma^m_{d]b}$ . We use the symbol ‘ $\equiv$ ’ to denote the fact that the equation defines a variable or a notation.

This review is structured in terms of five parts consisting of Sections 2 to 5. Section 2 introduces Lanczos-Lovelock models and their major structural and geometrical features. Section 3 begins by introducing the concept of horizon entropy in theories more general than Einstein gravity and uses it to study several aspects of horizon entropy in these models. Sections 4 and 5 describe the connection between Lanczos-Lovelock model and the emergent gravity paradigm.

## 2 Lanczos-Lovelock models: Structural Aspects

As emphasized above, one distinguishing feature of Lanczos-Lovelock models is that, although the action functional of the theory could be an arbitrary higher order polynomial in the curvature tensor, it leads to equations of motion which do not contain derivatives of the metric tensor of order higher than two. Originally, this class of theories were arrived at from a somewhat indirect route by Lovelock [2]. He wrote down the field equations in the form  $E_{ab} = (1/2)T_{ab}$  where  $T_{ab}$  is the energy momentum tensor of matter fields and demanded that  $E_{ab}$  satisfies the following requirements: (i)  $E_{ab}$  is symmetric in  $a$  and  $b$ ; (ii) it is a

function of metric and its first two derivatives; (iii) it is divergence-less. Having determined the most general form of  $E_{ab}$  satisfying these criteria, he obtained a Lagrangian from which one can obtain these field equations. (Note that we have not demanded  $E_{ab}$  to be *linear* in the second derivatives of the metric. If such a restriction is *also* added, one obtains Einstein's theory with a cosmological constant.) It is, however, possible to obtain Lanczos-Lovelock models in a more streamlined and transparent manner along different lines (see e.g., Chapter 15 of Ref. [6]) which is what we will pursue here.

## 2.1 Lagrangians from curvature and metric: general results

Let us consider a *general* Lagrangian built out of Riemann tensor and the metric, but not containing derivatives of the Riemann tensor, and ask the question: What is the condition on such a Lagrangian so as to ensure that equations of motion should contain no derivatives of the metric tensor higher than second order? The answer, as it turns out, uniquely picks out the Lanczos-Lovelock Lagrangians.

We will now describe this result highlighting the main conceptual points, many of which are of interest in their own right. We start with the action functional in  $D$ -dimensional spacetime:

$$A = \int_{\mathcal{V}} d^D x \sqrt{-g} L \quad (1)$$

where  $L$  is a scalar built from the metric and the curvature tensor. We can think of the Lagrangian  $L$  as a functional of any one of  $(R_{abcd}, R^a_{bcd}, R^{ab}_{cd} \dots)$  in combination with either  $g^{ab}$  or  $g_{ab}$ . The variation of the Lagrangian with respect to curvature and metric will bring in tensors defined by the derivative of Lagrangian with respect to different kinds of curvature components and the metric, like, for e.g.,

$$P^{abcd} \equiv \left( \frac{\partial L}{\partial R_{abcd}} \right)_{g_{ij}} ; \quad P^{ab} \equiv \left( \frac{\partial L}{\partial g_{ab}} \right)_{R_{ijkl}} \quad (2)$$

It is obvious from the above definition that  $P^{ab}$  is symmetric and  $P^{abcd}$  has the algebraic properties of the curvature tensor:

$$P^{abcd} = -P^{bacd} = -P^{abdc}, \quad P^{abcd} = P^{cdab}, \quad P^a{}^{[bcd]} = 0 \quad (3)$$

Curiously enough, the *mere fact that  $L$  is a scalar imposes several further relations* among these derivatives which play a crucial role in deriving the Lanczos-Lovelock field equations [7]. The most important ones among these relations can be expressed in terms of the tensor

$$\mathcal{R}^{ab} \equiv P^{aijk} R^b_{ijk} \quad (4)$$

which plays the role analogous to Ricci tensor of Einstein's theory. It can be shown (see [7]; Appendix A summarises these results with proof) that the following relations hold: To begin with,  $P^{ab}$  and  $\mathcal{R}^{ab}$  are closely related to each other:

$$\left( \frac{\partial L}{\partial g_{ab}} \right)_{R_{ijkl}} = -2R^b_{ijk} \left( \frac{\partial L}{\partial R_{aijk}} \right)_{g_{ab}} ; \quad \text{i.e., } P^{ab} = -2\mathcal{R}^{ab} \quad (5)$$

It immediately follows that  $\mathcal{R}^{ab}$  is symmetric; this is by no means an obvious result and does *not* follow from the algebraic properties of curvature tensor. Using Eq. (5), one can obtain several other relations connecting the partial derivatives when  $(g^{ab}, R_{abcd})$  or  $(g^{ab}, R^a_{bcd})$  or  $(g^{ab}, R^{ab}_{cd})$  are used as independent variables to describe the system. If we use the pair  $(g^{ab}, R_{abcd})$  as independent variables, we have

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} = - \left( \frac{\partial L}{\partial g_{im}} \right)_{R_{abcd}} = 2\mathcal{R}_{im} \quad (6)$$

On the other hand, if we use the pair  $(g^{ab}, R^a_{bcd})$  as independent variables, one can show that

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R^a_{bcd}} = \mathcal{R}_{im} \quad (7)$$

and, more interestingly, if we use the pair  $(g^{ab}, R^{ab}_{cd})$  as independent variables, we get

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R^{ab}_{cd}} = 0 \quad (8)$$

Finally, it can be proved that the tensor  $\nabla_m \nabla_n P^{amnb}$  is symmetric in  $a, b$ .

An *intuitive* way of understanding these results is as follows: To construct scalar *polynomials* of arbitrary degree in the curvature tensor, using only the curvature tensor  $R_{abcd}$  and metric  $g^{ij}$  (with index placements as chosen), we need two factors of the metric tensors to contract out the four indices of each curvature tensor in each term. So Eq. (6) is clearly true if  $L$  is a polynomial in  $R_{abcd}$ . Any arbitrary scalar function of curvature and metric can be represented as a (possibly infinite) series expansion in powers of the curvature tensor. Because Eq. (6) is linear in  $L$ , if it holds for a polynomial of arbitrary degree in curvature tensor, then it should hold for arbitrary scalar functions which possess a series expansion. This (intuitive) argument also gives the result in Eq. (8) which states that in scalar polynomials constructed from  $R^{ab}_{cd}$  the metric cannot appear and all the contractions must be with Kronecker deltas! This is obvious because, if we use  $g^{ab}$  to contract any two lower indices, that will leave two upper indices hanging loose which cannot be contracted out because we do not have  $g_{ab}$  available as an independent variable. Once we have Eq. (8) the other results can be obtained from it quite easily. Of course, the argument based on polynomials does not capture the full generality of our results, which hold for arbitrary Lagrangians that are not necessarily polynomials.

Having obtained these useful identities, let us now proceed with the variation of the action treating the Lagrangian as a function of  $R_{abcd}$  and  $g^{ab}$  with the index placements as shown. This leads to the result:

$$\delta A = \delta \int_{\mathcal{V}} d^D x \sqrt{-g} L = \int_{\mathcal{V}} d^D x \sqrt{-g} E_{ab} \delta g^{ab} + \int_{\mathcal{V}} d^D x \sqrt{-g} \nabla_j \delta v^j \quad (9)$$

where

$$\begin{aligned} E_{ab} &\equiv \frac{1}{\sqrt{-g}} \left( \frac{\partial \sqrt{-g} L}{\partial g^{ab}} \right)_{R_{abcd}} - \mathcal{R}_{ab} - 2\nabla^m \nabla^n P_{amnb} \\ &= \mathcal{R}_{ab} - \frac{1}{2} g_{ab} L - 2\nabla^m \nabla^n P_{amnb} \end{aligned} \quad (10)$$

and

$$\delta v^j \equiv [2P^{ibjd}(\nabla_b \delta g_{di}) - 2\delta g_{di}(\nabla_c P^{ijcd})] \quad (11)$$

In arriving at the second equality in Eq. (10) we have used Eq. (6). Note that each of the three terms in  $E_{ab}$  in Eq. (10) is individually symmetric in  $a, b$ .

In order to obtain sensible field equations from  $\delta A = 0$ , we need to ignore the boundary term involving  $\delta v^i$ . From Eq. (11) it is clear that both the variation of the metric and its first derivative normal to the boundary contributes at the boundary. In other words, we do *not* get a sensible variational principle if we only assume  $\delta g_{ab} = 0$  at the boundary. (This problem, of course, is well known even in Einstein's theory.) To obtain a well defined action principle, we have to add some boundary terms to our action such that the variations of the normal derivatives of the metric are cancelled out. We will say more about it in Sec. 2.6 but will assume, for the moment, that the boundary term involving  $\delta v^i$  can be ignored or cancelled. In that case, taking the total action to be  $A_{tot} = A + A_{matter}$  and defining the energy momentum tensor for matter from the variation of  $A_{matter}$  with respect to the metric in the usual manner, our field equations will reduce to  $E_{ab} = (1/2)T_{ab}$  where  $T_{ab}$  is the matter energy momentum tensor.

## 2.2 Lanczos-Lovelock action functional

Since  $P^{abcd}$  involves second derivatives of the metric, the term  $\nabla^m \nabla^n P_{amnb}$  in Eq. (10) will contain up to fourth derivatives of the metric. Hence, to satisfy our requirement that equations of motion should not have derivatives higher than second, we need to impose the condition:

$$\nabla_a P^{abcd} = 0 \quad (12)$$

(It turns out that the seemingly more general condition,  $\nabla^m \nabla^n P_{amnb} = 0$  does not lead to anything different.) Due to symmetries of  $P_{abcd}$ , this makes  $P^{abcd}$  divergence-free in all indices. In this case, the equations of motion become

$$\mathcal{R}_{ab} - \frac{1}{2}g_{ab}L = \left( \frac{\partial \sqrt{-g}L}{\partial g^{ab}} \right)_{R^a_{bcd}} = \frac{1}{2}T_{ab} \quad (13)$$

In arriving at the first equality we have used Eq. (7). Clearly, the field equations do not involve more than second derivatives of the metric but are, in general, nonlinear functions of the second derivatives (the only exception being the Einstein-Hilbert Lagrangian). Note that the structure of  $E_{ab}$  is very similar to that in Einstein's theory with  $\mathcal{R}_{ab}$  playing the role analogous to Ricci tensor.

Our original task of finding an action functional, leading to equations of motion which contain no derivatives higher than the second order, now reduces to finding all scalar functions of curvature and metric such that Eq. (12) is satisfied. We shall now describe how one such class of Lagrangians can be constructed but will not bother to prove that our construction is actually unique — for which we refer the reader to original literature [1, 2].

For this purpose, let us consider a general class of Lagrangians of polynomial form in curvature:

$$L = \sum_{m=0}^M L_m [(\text{Riem})^m] \quad (14)$$

where  $L_m$  contains  $m$  factors of the Riemann tensor, chosen with index placements  $R_{cd}^{ab}$ . Our condition in Eq. (8) then implies that the Lagrangian is independent of the metric tensor and all the indices of the curvature tensor  $R_{cd}^{ab}$  must be contracted out using Kronecker deltas. All we have to do is to find the correct manner of contracting the indices so as to satisfy Eq. (12). Ignoring a constant term corresponding to  $m = 0$ , the next simplest case corresponds to  $m = 1$  which should be a Lagrangian linear in  $R_{cd}^{ab}$ . The only way of contracting the indices of  $R_{cd}^{ab}$  using Kronecker deltas leads to

$$L_{EH} = \delta_{24}^{13} R_{13}^{24}, \quad \delta_{24}^{13} = \delta_2^1 \delta_4^3 - \delta_2^3 \delta_4^1 \quad (15)$$

where  $\delta_{24}^{13}$  is the determinant tensor involving products of two Kronecker deltas and we are using an obvious notation as regards contraction of indices where the numeral  $n$  actually stands for an index  $a_n$  etc. This clearly leads to Einstein's theory with  $L_{EH} \propto R$ .

The next order Lagrangian (for  $m = 2$ ) will be quadratic in  $R_{cd}^{ab}$ . This implies that  $P^{abcd}$  (defined by Eq. (2)) will be linear in the curvature tensor. The corresponding Lagrangian can be found by obtaining the most general fourth rank tensor built from metric and curvature which is (i) linear in curvature and (ii) possesses symmetries of curvature tensor and (iii) divergence-free in all indices. The choice is unique and given by

$$P^{abcd} = R^{abcd} - G^{ac} g^{bd} + G^{bc} g^{ad} + R^{ad} g^{bc} - R^{bd} g^{ac} \quad (16)$$

(In four dimensions, this tensor is essentially the double-dual of Riemann.) Using this we can find the  $m = 2$  Lagrangian to be

$$L = \frac{1}{2} (g_{ia} g^{bj} g^{ck} g^{dl} - 4g_{ia} g^{bd} g^{ck} g^{jl} + \delta_a^c \delta_i^k g^{bd} g^{jl}) R^i_{jkl} R^a_{bcd} = \frac{1}{2} [R^{abcd} R_{abcd} - 4R^{ab} R_{ab} + R^2] \quad (17)$$

This is the first non-trivial example of a Lagrangian leading to second order equations of motion from a Lagrangian which is nonlinear in the curvature tensor. The Lagrangian above is called Gauss-Bonnet Lagrangian, and a more explicit construction of it is given in Appendix B. The Gauss-Bonnet Lagrangian can also be written in the form:

$$L_{GB} = \delta_{2468}^{1357} R_{13}^{24} R_{57}^{68} \quad (18)$$

where  $\delta_{2468}^{1357}$  is again the determinant tensor involving four factors of Kronecker deltas.

The examples of Einstein-Hilbert and Gauss-Bonnet actions suggest an obvious generalization for higher orders that indeed leads to the Lanczos-Lovelock Lagrangian, which — in  $D$ -dimensions — is given by the sum:

$$L = \sum_m c_m L_m \quad (19)$$

$$L_m^{(D)} = \frac{1}{16\pi} \frac{1}{2^m} \delta_{c_1 d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} \quad (20)$$

where the tensor appearing in the right hand side of Eq. (20) is a completely antisymmetric

determinant tensor defined as:

$$\delta_{jc_1d_1\dots c_md_m}^{ia_1b_1\dots a_mb_m} = \det \left[ \begin{array}{c|cccc} \delta_j^i & \delta_{c_1}^i & \dots & \delta_{d_m}^i \\ \hline \delta_j^{a_1} & & & \\ \vdots & & \delta_{c_1d_1\dots c_md_m}^{a_1b_1\dots a_mb_m} & \\ \delta_j^{b_m} & & & \end{array} \right] \quad (21)$$

for  $m \geq 0$ . The lowest order terms,  $m = 0, 1$  correspond to cosmological constant and the Einstein-Hilbert action respectively, as can be easily seen by expanding the alternating determinant. For  $m = 1$ ,  $L_1 = (16\pi)^{-1}R$ , and the factor of  $16\pi$  in the definition of  $L_m$  essentially changes the right hand side of equations of motion from the conventional  $8\pi T_{ab}$  to  $(1/2)T_{ab}$ . (Getting an *explicit* expression for  $L_m$ 's remains a complicated task since it requires evaluating the determinant of a  $2m \times 2m$  matrix; fortunately, we shall never require explicit form of  $L_m$ 's for the results which we wish to cover in this review!)

The  $c_m$ 's in Eq. (19) are arbitrary coupling constants which are a priori unconstrained. Some restrictions on  $c_m$ 's might arise while considering specific solutions in Lanczos-Lovelock theory (such as black holes), or these might be related to certain parameters of a fundamental (quantum) theory when we consider low energy effective action that contain such Lanczos-Lovelock terms. This second situation arises in string theory where  $c_2$ , for e.g., is related to the string scale,  $c_2 \propto l_s^2$ . For most of our discussion in this review, we shall assume  $c_m$ 's to be arbitrary.

To complete our argument we now need to prove that, for this Lagrangian, Eq. (12) is indeed satisfied. To do this, note that, being homogeneous in  $R_{abcd}$ , the  $L_m$ 's can be written as

$$L_m = \frac{1}{m} \frac{\partial L_m}{\partial R_{abcd}} R_{abcd} = (1/m) P_{(m)}^{abcd} R_{abcd} \quad (22)$$

It is obvious that this tensor is same as the tensor  $P_{abcd}$  defined earlier, restricted to a particular value of  $m$ . The condition in Eq. (12) is therefore ensured if

$$\nabla_a P_{(m)}^{abcd} = 0 \quad (23)$$

This is easily proved using the standard Bianchi identity for the curvature tensor:

$$\nabla_a P_{cd(m)}^{ab} \propto \nabla_a \delta_{cd\star\star\dots\star}^{ab\star\star\dots\star} \underbrace{R_{\star\star}^{\star\star} \dots R_{\star\star}^{\star\star}}_{(m-1)\text{factors}} \quad (24)$$

$$\propto \delta_{cd\star\star\dots\star}^{ab\star\star\dots\star} \nabla_{[a} R_{\star\star]}^{\star\star} \dots R_{\star\star}^{\star\star} \quad (25)$$

$$= 0 \quad (26)$$

where in the second step we have used complete antisymmetry of determinant tensor, and the final equality follows from the Bianchi identity  $\nabla_{[a} R_{bc]}^{ij} = 0$ . We have also introduced a convenient symbolic notation of using similar symbols (such as  $\star, \bullet$  etc.) to denote contracted indices. This helps us display relevant expressions (especially in later sections) in a form which



conveys same information in a much more transparent manner. Therefore, the Lanczos-Lovelock Lagrangians certainly would lead to desired equations of motion (not containing higher derivatives of the Riemann tensor). That they are unique in doing so requires some more work, but the demonstration for Gauss-Bonnet given in Appendix B should at least make it plausible.

Just as in the above result, almost all of the unique features of the Lanczos-Lovelock models arise due to (i) the combinatorial properties of the determinant tensor and (ii) the fact that the  $L_m$ 's are homogeneous polynomials of order  $m$  in the Riemann tensor. In particular, the Lanczos-Lovelock Lagrangians are *finite* order polynomials; i.e., although the sum in Eq. (19) is defined for all  $m \in \mathbb{Z}$ , only a finite number of  $L_m$ 's are non-zero in a given spacetime dimension  $D$ . This follows essentially from complete antisymmetry of the determinant tensor and can be seen as follows. For even  $D$ , the determinant tensor vanishes if  $m > D/2$ , whereas for odd  $D$ , it vanishes if  $m > (D-1)/2$ . Therefore, for a given  $D$ , there is a critical value  $m = m_c$  such that the action is non-zero only for  $L_m$ 's with  $m \leq m_c$ , where  $m_c = [(D-1)/2]$  with  $[x]$  denoting the smallest integer not smaller than  $x$ . The dimension  $D = 2m$  is called the *critical dimension* for a given  $L_m^D$ , since  $L_m^{D < 2m}$  vanishes identically.

The structure of the Lanczos-Lovelock action in  $D = 2m$  is interesting in its own right, and highlights certain topological significance of the Lanczos-Lovelock actions. This will be discussed in detail in Sec. 2.5. In particular, the critical dimension for Einstein-Hilbert action ( $m = 1$ ) is  $D = 2$  in which  $G_{ab}$  identically vanishes; similarly the critical dimension for the Gauss-Bonnet Lagrangian ( $m = 2$ ) is  $D = 4$  and it does not contribute to the equations of motion in  $D = 4$ . We will say more about critical dimensions in Sec. 2.5.

It is possible to write down an alternate form of the Lanczos-Lovelock action, which will be shown (in Sec. 2.7) to be related to the 'holographic' structure of these actions. This alternate form is a direct generalization of the so called  $\Gamma\Gamma$  (or Dirac) Lagrangian for Einstein gravity. We will sketch the basic structure symbolically, and will describe the details while discussing holographic structure of Lanczos-Lovelock actions in Sec. 2.5. Consider any Lagrangian which can be written as  $L = Q_{abcd}R^{abcd}$  such that (i)  $\nabla_a Q^{abcd} = 0$  and (ii)  $Q_{abcd}$  has all the algebraic symmetries of Riemann tensor. We know from the above discussion that Lanczos-Lovelock models belong to this class. We can then (symbolically) decompose  $L$  as follows

$$\begin{aligned} L \sim QR &\sim Q(\partial\Gamma + \Gamma\Gamma) \\ &\sim Q\Gamma\Gamma + \partial(Q\Gamma) + \Gamma\nabla Q \\ &\sim Q\Gamma\Gamma + \partial(Q\Gamma) \end{aligned} \tag{27}$$

One immediately sees, in the final expression above, (a generalization of) the structure of the  $\Gamma\Gamma$  Lagrangian familiar from Einstein theory. The surface term in the above split is important in studying the thermodynamical properties of horizons in Lanczos-Lovelock models, as we shall show in a later section.

Before concluding this section we will also comment on the Palatini form of the variational principle for the general Lagrangian. In Einstein theory, the conventional metric variation of the action yield equations of motion which are equivalent to those obtained upon treating metric and connection as independent quantities and varying  $g^{ab}$  and  $\Gamma^i_{jk}$  independently. The latter variational principle (the so called Palatini variation) actually forces the connection to be metric compatible. For the general Lagrangian  $L[g^{ab}, R_{abcd}]$ , it is easy to show that, when

the connection and metric are varied independently, we obtain

$$\delta(L_m \sqrt{-g}) = \nabla_i \left( 2\sqrt{-g} g^{al} P_{kl}^{ij} \delta \hat{\Gamma}_{ja}^k \right) + \sqrt{-g} [\mathcal{R}_{ab} - \frac{1}{2} g_{ab} L] \delta g^{ab} - \nabla_i \left( 2P_{kl}^{ij} g^{al} \sqrt{-g} \right) \delta \hat{\Gamma}_{ja}^k \quad (28)$$

where  $\hat{\Gamma}_{ja}^k$  is an independent connection. If we ignore the total divergence term and set the coefficients of  $\delta g^{ab}$  and  $\delta \hat{\Gamma}_{ja}^k$  independently to zero, then we get the conditions

$$\mathcal{R}_{ab} - \frac{1}{2} g_{ab} L = \frac{1}{2} T_{ab}; \quad \nabla_i \left( 2P_{kl}^{ij} g^{al} \sqrt{-g} \right) = 0 \quad (29)$$

It is difficult to make sense of the second equation in general and obtain the correct field equations. But if we restrict ourselves to Lagrangians which satisfy the condition  $\nabla_i P_{kl}^{ij} = 0$ , then the vanishing of the coefficient of  $\delta \hat{\Gamma}_{ja}^k$  requires  $\nabla_i (g^{al} \sqrt{-g}) = 0$ . This in turn can be shown to lead to  $\nabla_i \sqrt{-g} = 0 = \nabla_i g_{ab}$ , leading to the standard relation between metric and connection. The vanishing of the coefficient of  $\delta g^{ab}$  now leads to the correct equations of motion. Thus for the Lanczos-Lovelock Lagrangians, the Palatini form of the variation leads to the same field equations, just as in Einstein's theory.

The converse is also true [8]. If we assume that  $\nabla_i \sqrt{-g} = 0 = \nabla_i g_{ab}$  (which makes connection and metric compatible), then the second condition in Eq. (29) leads to  $\nabla_i P_{kl}^{ij} = 0$ . That is, the compatibility of metric and connection (which is sometimes taken as a manifestation of the equivalence principle) for the generic Lagrangian of the form  $L[g_{ab}, R_{abcd}]$  picks out Lanczos-Lovelock Lagrangians uniquely. (Same conclusion was also reached in [9]. See also [10].)

### 2.3 Explicit form of the equations of motion

The equations of motion for the Lanczos-Lovelock model, given by Eq. (13), can be explicitly computed by first determining the form of  $P^{abcd}$  from the Lagrangian in Eq. (20). This gives, for a generic Lanczos-Lovelock action  $L = \sum_m c_m L_m$ , the equations:

$$E_b^a = \sum_m c_m E_{b(m)}^a = \frac{1}{2} T_b^a$$

where

$$\begin{aligned} E_{j(m)}^i &= \frac{1}{16\pi} \frac{m}{2^m} \delta_j^{a_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{i d_1} \dots R_{a_m b_m}^{c_m d_m} - \frac{1}{2} \delta_j^i L_m \\ &= -\frac{1}{2} \frac{1}{16\pi} \frac{1}{2^m} \delta_{j c_1 d_1 \dots c_m d_m}^{i a_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} \end{aligned} \quad (30)$$

The equivalence of the two expressions in Eq. (30) can be established through the identity;

$$\left[ \delta_j^i \delta_{c_1 d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} - 2m \delta_{c_1}^i \delta_{j d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} \right] R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} = \delta_{j c_1 d_1 \dots c_m d_m}^{i a_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} \quad (31)$$

which, in turn, is easily verified by noting that the alternating tensor in the right hand side of Eq. (31) can be written as a determinant:

$$\delta_{jc_1 d_1 \dots c_m d_m}^{ia_1 b_1 \dots a_m b_m} = \det \left[ \begin{array}{c|cccc} \delta_j^i & \delta_{c_1}^i & \dots & \delta_{d_m}^i \\ \hline \delta_j^{a_1} & & & \\ \vdots & & \delta_{c_1 d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} & \\ \delta_j^{b_m} & & & \end{array} \right] \quad (32)$$

The first term on the left-hand-side of Eq. (31) therefore comes from the multiplication of  $\delta_j^i$  with the lower right block of the above matrix. The remaining terms in the determinant can be grouped as

$$- \left[ \delta_{c_k}^i \delta_{jc_1 d_1 \dots d_{k-1} d_k \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} - \delta_{d_k}^i \delta_{jc_1 d_1 \dots c_k c_{k+1} \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} \right] \quad (33)$$

where  $1 \leq k \leq m$ . Since this whole determinant is multiplied by the product of curvature tensors, only the piece antisymmetric in the pair  $\{c_k, d_k\}$  will be picked up, producing a factor of 2 for each pair  $\{c_k, d_k\}$ . Further, each of the  $m$  such pairs contribute the same amount due to the symmetries of the alternating tensor and the curvature tensor. This gives another factor of  $m$ , so that the contribution of the remaining terms in the above determinant becomes equal to  $2m$  times the contribution of any particular term, say, the term corresponding to the pair  $\{c_1, d_1\}$ . Noting the overall minus sign in (33), we obtain the second term on the left-hand-side of Eq. (31), thereby proving the desired equality.

There are three more results which are easily obtained from the explicit form of the equations of motion:

1. The field equations resulting from a  $m$ -th order Lanczos-Lovelock Lagrangian, which involves  $E_b^a$  in Eq. (30), become vacuous when  $D = 2m$ , the critical dimension for a given Lanczos-Lovelock term. This is immediately obvious from the second form of Eq. (30), on recalling that the determinant tensor is completely anti-symmetric, and hence the number of values that the  $2m + 1$  indices of the tensor can take (which of course is the same as the number of spacetime dimensions  $D$ ), must be greater than or equal to  $2m + 1$ . For  $D \leq 2m$  the  $E_b^a$  vanishes identically.
2. Since  $E_{ab} = \mathcal{R}_{ab} - \frac{1}{2}g_{ab}L$  arises from the variation of a generally covariant scalar with respect to  $g^{ab}$ , it follows that  $\nabla_a E_b^a = 0$  which is the analogue of standard Bianchi identity. (The proof proceeds exactly as in the case of standard Bianchi identity; see e.g., page 255 of Ref. [6]). This in turn implies that

$$\nabla_a \left( P_b^{kij} R_{kij}^a \right) = \frac{1}{2} \partial_b L \quad (34)$$

One consequence of this result is the following: Suppose we have a theory in which the field equations given by are the projection of the Lanczos-Lovelock field equations on to null vectors. That is, suppose the field equations of the theory are given by:

$$2P_b^{kij} R_{kij}^a n_a = T_b^a n^b n_a \quad (35)$$

for *all* null vectors  $n_a$  with  $T_b^a$  being a covariantly conserved symmetric tensor. This implies the equations

$$2E_b^a = \left[ 2P_b^{ijk} R_{ijk}^a - \delta_b^a L \right] = T_b^a + \Lambda \delta_b^a \quad (36)$$

where  $\Lambda$  is a *constant*. To see this, note that Eq. (35) implies that  $2P_b^{ijk} R_{ijk}^a - T_b^a = -\lambda \delta_b^a$  where  $\lambda(x)$  is some scalar. Rewriting this equation as  $2E_b^a + (L + \lambda) \delta_b^a = T_b^a$ , taking the divergence and using  $\nabla_a E_b^a = 0 = \nabla_a T_b^a$  we get  $\partial_b(L + \lambda) = 0$ . Therefore,  $\lambda = -L + \text{constant}$ . In other words, a theory with null projection of the equations of motion is equivalent to the original theory with a cosmological constant added to it.

3. It can be shown that [11] for each Lanczos-Lovelock term of order  $m$ , one can construct a fourth rank tensor  $\mathcal{R}_{abcd}^{(m)}$  from  $P_{abcd}^{(m)}$  and  $R_{abcd}^{(m)}$ , with all the algebraic symmetries of the Riemann tensor and having the property that, the trace of it's Bianchi derivative  $\nabla_{[a} \mathcal{R}_{|j|bc]}^{i(m)}$  vanishes, yielding a divergence-free tensor which is the analogue of Einstein tensor for Lanczos-Lovelock models. Moreover, it can be argued [11] that Lanczos-Lovelock Lagrangians are the only ones which admit such a description in terms of a Bianchi derivative. It is also worth mentioning that, in [12], it has been proved, using the above mentioned generalized Riemann tensor for pure Lanczos-Lovelock actions, that for any pure Lanczos-Lovelock term of order  $m$ , the vacuum is trivial with respect to  $\mathcal{R}_{jkl}^{i(m)}$  in all odd dimensions  $D = 2m + 1$ ; i.e., it corresponds to  $\mathcal{R}_{jkl}^{i(m)} = 0$ , which one might call as *Lovelock-flat*. This generalizes the well known fact in Einstein gravity ( $m = 1$ ), that the vacuum solutions in  $D = 3$  are Riemann flat.

The additional characterizations of Lanczos-Lovelock models like the one mentioned above (as well as the one arising from validity of the Palatini variational principle) are not connected (at least in any obvious way) to the original requirement of quasi-linearity of equations of motion from which these models were first obtained. While it is possible that all such “uniqueness” results for Lanczos-Lovelock theories are eventually connected to (and derivable from) the topological significance of the Lanczos-Lovelock action in the critical dimension, we do not have at present in the literature a proof of the same. Such a proof might have important implications for several of the results we shall discuss in this review, which hinge on very special algebraic properties of certain tensors encountered in the Lanczos-Lovelock theory.

4. It also follows from Eq. (13) that, for the  $m$ th Lanczos-Lovelock Lagrangian,  $L_{(m)}$ , the trace of the equations of motion is proportional to the Lagrangian:

$$E \equiv g^{ab} E_{ab} = \frac{g^{ab}}{\sqrt{-g}} \frac{\partial \sqrt{-g} L_{(m)}}{\partial g^{ab}} = -[(D/2) - m] L_{(m)} \quad (37)$$

This *off-shell* relation (that is, a relation valid even field equations are not imposed) is easy to prove from the fact that we need to introduce  $m$  factors of  $g^{ab}$  to proceed from  $R_{bcd}^a$  to  $R_{cd}^{ab}$  and that  $\sqrt{-g}$  is homogeneous function of  $g^{ab}$  of degree  $-D/2$ . Using this we get  $L_{(m)} = -2E/(D - 2m) = -T/(D - 2m)$  which allows us to rewrite the field equations of the  $m$ -th order Lanczos-Lovelock theory in the form:

$$2P_b^{ijk} R_{ijk}^a = T_b^a - \delta_b^a \frac{1}{(D - 2m)} T \quad (38)$$

This is similar in structure to  $R_b^a = 8\pi(T_b^a - \delta_b^a(1/2)T)$ , to which it reduces to in  $D = 4$  for the  $m = 1$  case. Note that this result does *not* hold for the theory obtained by adding together Lanczos-Lovelock terms of different order.

## 2.4 Diffeomorphism invariance and the Noether current

In any generally covariant theory, the infinitesimal coordinate transformations  $x^a \rightarrow x^a + \xi^a(x)$  lead to the conservation of a current (called the Noether current) that can be obtained as follows: The variation of the gravitational Lagrangian resulting from arbitrary variations of  $\delta g^{ab}$  can be expressed as form,

$$\delta(L\sqrt{-g}) = \sqrt{-g} (E_{ab}\delta g^{ab} + \nabla_a \delta v^a). \quad (39)$$

where the  $\delta v^a$  term leads to the boundary contribution. When the variations in  $\delta g^{ab}$  arise due to the diffeomorphism  $x^a \rightarrow x^a + \xi^a$  we have,  $\delta(L\sqrt{-g}) = -\sqrt{-g}\nabla_a(L\xi^a)$ , with  $\delta g^{ab} = (\nabla^a \xi^b + \nabla^b \xi^a)$ . Substituting these in Eq. (39) and using the (generalized) Bianchi identity  $\nabla_a E^{ab} = 0$ , we obtain the off-shell conservation law  $\nabla_a J^a = 0$ , for the current,

$$J^a \equiv (2E^{ab}\xi_b + L\xi^a + \delta_\xi v^a) \quad (40)$$

where  $\delta_\xi v^a$  represents the boundary term which arises for the specific variation of the metric in the form  $\delta g^{ab} = (\nabla^a \xi^b + \nabla^b \xi^a)$ . It is also convenient to introduce the antisymmetric tensor (called the Noether potential)  $J^{ab}$  by  $J^a \equiv \nabla_b J^{ab}$ . Using the known expression in Eq. (11) for  $\delta_\xi v^a$  in Eq. (40), it is possible to write an explicit expression for the current  $J^a$  for any diffeomorphism invariant theory. For the general class of theories we are considering, the  $J^{ab}$  and  $J^a$  can be expressed in the form

$$J^{ab} = 2P^{abcd}\nabla_c \xi_d - 4\xi_d (\nabla_c P^{abcd}) \quad (41)$$

$$J^a = -2\nabla_b (P^{ab}{}^c + P^{ac}{}^b) \nabla_c \xi_d + 2P^{abcd}\nabla_b \nabla_c \xi_d - 4\xi_d \nabla_b \nabla_c P^{abcd} \quad (42)$$

where  $P_{abcd} \equiv (\partial L / \partial R^{abcd})$ . The Lagrangians we are considering here depend, upto this point, on metric and Riemann tensor, but *not* on higher derivatives of the curvature tensor. However, the expressions in this section hold even in the latter case, the only difference being that the tensor  $P_{abcd}$  will now be defined with additional terms involving terms such as  $\partial L / \partial (\nabla R)$  etc. Algebraically,  $P_{abcd}$  simply gives the “equation of motion” with  $R_{abcd}$  treated as the dynamical variable.

In the case of Lanczos-Lovelock models,  $\nabla_a P^{abcd} = 0$  and one can obtain from the Eq. (11) the expression for the boundary term  $\delta v^i$  to be:

$$\delta v^i = -2P_k{}^{aj} \delta_\xi \Gamma_{aj}^k \quad (43)$$

where the Lie derivative of the connection (which is a generally covariant tensor) is defined through:

$$\mathcal{L}_\xi \Gamma_{ab}^c = \nabla_a \nabla_b \xi^c - R_{abm}^c \xi^m \quad (44)$$

This leads to the expressions for Noether potential and current given by

$$J^{ab} = 2P^{abcd}\nabla_c \xi_d; \quad J^a = 2P^{abcd}\nabla_b \nabla_c \xi_d = (2\mathcal{R}^{ab}\xi_b + \delta_\xi v^a) \quad (45)$$

In this case, we can also write the current as:

$$J^i = 2\mathcal{R}^{ib}\xi_b + \delta_\xi v^i = 2\mathcal{R}^{ab}\xi_b - 2P_k^{aji} \delta_\xi \Gamma_{aj}^k \quad (46)$$

This result shows that, near any event  $\mathcal{P}$  where  $\xi^a$  behaves like an (approximate) Killing vector and satisfies the conditions  $\mathcal{L}_\xi g_{ab} = \nabla_{(a}\xi_{b)} = 0$  and  $\mathcal{L}_\xi \Gamma_{ab}^c = 0$  (which a true Killing vector will satisfy everywhere), we get

$$J^a \equiv 2\mathcal{R}^{ab}\xi_b. \quad (47)$$

Incidentally, the expression for the Noether current can also be obtained from the Palatini variation in Eq. (28). Assuming that the metric is compatible with the connection and considering variations arising from diffeomorphisms for which  $\delta(L_m\sqrt{-g}) = -\sqrt{-g} \nabla_i (L\xi^i)$ , we obtain an expression for the conserved current as:

$$\nabla_i \left( 2E_j^i \xi^j + L\xi^i + 2g^{al} P_{kl}^{ij} \delta_\xi \Gamma_{ja}^k \right) = 0 \quad (48)$$

where  $\Gamma_{bc}^a$  is now the metric compatible connection. This yields the same expression as in Eq. (46).

## 2.5 Action functional in the critical dimension

As we said before, the field equations resulting from a  $m$ -th order Lanczos-Lovelock Lagrangian become trivial when  $D = 2m$ , known as the critical dimension for a given Lanczos-Lovelock term. The Lanczos-Lovelock action itself also vanishes identically for  $D < 2m$ . But for  $D = 2m$ , the action itself does not vanish. This raises the question: What is the structure of Lanczos-Lovelock action itself in the critical dimension? The vacuous nature of field equations requires that the *variation* of the action will be a pure surface term in the critical dimension. But it is not obvious whether the *action itself* is a pure surface term. We discuss these and related issues in this section, beginning by stressing why such issues are interesting and non-trivial.

Let us begin with  $D = 2$ , which is the critical dimension for  $m = 1$  Lanczos-Lovelock theory, the Lagrangian of which is simply  $R$  leading to Einstein-Hilbert action. It is trivial to show, using the language of differential forms, that the integrand of the Einstein-Hilbert action  $R\sqrt{-g}$  in  $D = 2$  can be written as an exterior derivative of a quantity which can be expressed in terms of the *tetrads* (see Eq. (50) below). But what we are interested in is whether or not  $R\sqrt{-g}$  can be expressed as  $\partial_j R^j$  for a doublet of functions  $(R^0, R^1)$  built from the *metric and its derivatives* without using tetrads in a *general* coordinate system. That is, the line element is:

$$ds^2 = A(t, x)dt^2 + 2C(t, x)dxdt + B(t, x)dx^2, \quad (49)$$

(in which we have not imposed any gauge or coordinate conditions) we want to find an explicit expression for  $(R^0, R^1)$  in terms of  $A, B, C$  and their derivatives. This is again trivial to accomplish in the conformally flat gauge with  $C = 0, A = B$  in which  $R\sqrt{-g} = \partial_j \partial^j \ln A$ . This has the required form but it will *not* survive if we transform to an arbitrary coordinate system.

Since we already know the answer in terms of differential forms, one might think that there is no room for controversy about obtaining an explicit expression for  $R^j$ . However, a paper

by Deser and Jackiw [13] seems to suggest that this is impossible, which led Kiriushcheva and Kuzmin [14] to conclude that there are fundamental differences between the vierbein formalism and the standard formalism and, because of those differences, the Einstein-Hilbert Lagrangian being a total derivative in the vierbein formalism does not imply that it is a total derivative in the standard formalism. Obviously, this issue deserves clarification: on one hand, several papers claim (at least implicitly) that the action itself is a total derivative for all Lanczos-Lovelock models in critical dimensions but without concrete proof; on the other hand, there are arguments, summarized in the articles by Kiriushcheva and Kuzmin [15, 14], that these claims are incorrect.

These contradictory conclusions originate largely due to lack of an explicit expression for  $R^j$ , as a function of the metric and its derivatives, satisfying  $L_m \sqrt{-g} = \partial_j R^j$ . We will first summarise some recent results [16] establishing that  $R^j$  can indeed be found in terms of the metric and its derivatives and giving an explicit expression for  $R^j$  for the case of  $D = 2$ . This  $R^j$  is not unique since one can always add to it any set of functions  $f^j(x)$  which satisfy  $\partial_j f^j = 0$ . There is, in fact, one natural  $R^i$  associated with each unit normalized vector field in  $D = 2$  and any two of them differ by an  $f^j$  with  $\partial_j f^j = 0$ .

We will also summarize similar results for  $D = 4$ . The Lanczos-Lovelock term, for which  $D = 4$  is the critical dimension, is the  $m = 2$  Gauss-Bonnet term. The question again is whether one can write  $L_{\text{GB}} \sqrt{-g} = \partial_j R^j$ . An expression for  $R^j$  was given in [17] based on an expression originally given in [18]. Unfortunately, this expression was obtained through an invalid identification of frame indices with spacetime indices and turns out to be incorrect. We will describe how correct expressions can be obtained and outline a proof that any Lanczos-Lovelock Lagrangian in the critical-dimension can be written as a total derivative of functions of the metric and its derivatives.

Let us begin with  $D = 2$ . In reference [16], three distinct methods for obtaining an explicit expression for  $R^j$  in this case are given, of which we discuss the one based on Cartan formalism because it is easy to generalise to an arbitrary Lanczos-Lovelock term,  $L_m$ . For this purpose, we introduce an orthonormal basis  $\omega^a = \omega_i^a dx^i$  (satisfying  $g_{ij} = \omega_i^a \omega_j^b \eta_{ab}$ ) as a basis for the cotangent space. The spin connection, defined by the first Cartan structural equation (see, for e.g., Appendix A of [16]) must be antisymmetric:  $\Omega_{01} = -\Omega_{10}$ . The Einstein-Hilbert two-form can be written (up to an overall sign) as:

$$\begin{aligned} R\sqrt{-g}d^2x &= \Theta^{ab} \wedge *(\omega_a \wedge \omega_b) \\ &= d\Omega_{01}d^2x, \end{aligned} \tag{50}$$

where  $\Theta^{ab} = (d\Omega^{ab} + \Omega_s^a \wedge \Omega^{sb})$  is the curvature two-form. So, to find  $R^j$  such that  $R\sqrt{-g} = \partial_j R^j$  we only need to calculate explicitly the spin connection  $\Omega_{01}$  and take its exterior derivative. This gives [16] the result:

$$\begin{aligned} R^0 &= \frac{1}{\sqrt{-g}} \left[ -\lambda \frac{g_{01}}{g_{00}} g_{00,1} + (1-\lambda) \frac{g_{01}}{g_{11}} g_{11,1} - 2(1-\lambda) g_{01,1} + g_{11,0} \right] \\ R^1 &= \frac{1}{\sqrt{-g}} \left[ -(1-\lambda) \frac{g_{01}}{g_{11}} g_{11,0} + \lambda \frac{g_{01}}{g_{00}} g_{00,0} - 2\lambda g_{01,0} + g_{00,1} \right], \end{aligned} \tag{51}$$

which is an one-parameter family of solutions labelled by a constant  $\lambda$ . This parameter can be thought of as describing the gauge freedom available in defining an orthonormal basis for



our cotangent space. The  $R^j$ 's for two different values of  $\lambda$  are an example of non-uniqueness mentioned earlier. One can directly verify that  $f^j \equiv [R^j(\lambda_1) - R^j(\lambda_2)]$  satisfies  $\partial_j f^j = 0$ , as it should. This gauge freedom is entirely unphysical, and one can fix it without any extra condition on the spacetime or the metric. The gauge-fixed Cartan formalism is equivalent to the standard formalism for General Relativity. An explicit expression for  $R^j$  for the  $m = 2$  (the Gauss-Bonnet term) can also be found in Ref. [16].

We will now outline a general method for finding  $R^j$ 's for a generic Lanczos-Lovelock term, referring the reader to [16] for technical details. There exists a simple geometrical interpretation for the Lanczos-Lovelock Lagrangians being total derivatives of expressions involving tetrads. The Lagrangian  $L_m$  is simply the Euler density of a  $2m$ -dimensional manifold (without a boundary), as can be seen, for example, using the results in the section on index theorems in ref. [19]. In  $2m$  dimensions,  $L_m$  is a topologically invariant scalar density because it is a closed form and therefore locally exact [20]. The Chern-Simons form  $Q_m$  is defined to quantify this equality: evaluating it between two connections  $\Omega$  and  $\Omega'$ , one obtains  $L_m(\Omega) - L_m(\Omega') = dQ_m(\Omega, \Omega')$ . In particular, if we restrict ourselves to a local coordinate patch, then we can introduce a flat connection  $\Omega' = 0$  and can therefore write  $L_m = dQ_m(\Omega, 0)$ . Hence, the existence of the Chern-Simons form is a valid proof that  $L_m$  is an exact differential. Moreover, this means that  $R^j$  can be thought of as providing the local coordinate description of the Chern-Simons form. Once it is clear that we can express  $L_m$  as an exact differential in the form language, we can proceed as in the two-dimensional case, and compute the set of (non-unique) functions  $R^j$  such that  $L_m \sqrt{-g} = \partial_j R^j$ . One simply has to define an orthonormal basis  $\omega^a$ , calculate the spin connections  $\Omega^{ab}$ , and use them in the expression already known in terms of differential forms. Therefore,  $L_m \sqrt{-g}$  can be written as a total derivative of a set of functions of the metric and its derivatives in any critical dimension.

We conclude this section by showing how  $R^j$  takes a very simple form in any spacetime possessing a Killing vector. This is a direct consequence of the results obtained in ref.[21] and the explicit proof goes as follows: Suppose the spacetime metric admits a timelike Killing vector  $\xi^a$  which we can choose, without loss of generality, to be:  $\xi^a = (1, 0, \dots, 0)$ . Since  $E_{ab} = 0$ , in the critical dimension, it follows from Eq. (10) in the case of Lanczos-Lovelock models with  $\nabla_a P^{abcd} = 0$  that

$$L_m = 2P^{0cde} R_{0cde} \quad (52)$$

We now note that

$$L_m = 2P^{0cdb} R_{0cdb} = P^{0cdb} R_{acdb} \xi^a = 2P^{0cdb} \nabla_c \nabla_d \xi_b = \frac{1}{\sqrt{-g}} \partial_c (2\sqrt{-g} P^{0cdb} \nabla_d \xi_b) \quad (53)$$

where we used the facts that  $P^{abcd}$  is divergence-free and has the same symmetries as the Riemann tensor. It follows that  $L_m \sqrt{-g} = \partial_c R^c$  with  $R^c = 2\sqrt{-g} P^{0cdb} \nabla_d \xi_b$ . Using  $\nabla_d \xi_b = \Gamma_{b0d}$ , we find that

$$R^j = -2\sqrt{-g} P_a{}^{bj0} \Gamma_{0b}^a \quad (54)$$

In this case  $R^j$  is actually a component of the Noether potential corresponding to diffeomorphism invariance, giving a direct physical interpretation to the local version of the Chern-Simons form in spacetimes which have symmetries. (For such spacetimes, the Noether current also gives the horizon entropy provided a Killing horizon exists; we shall discuss this further in Sec. 3.1.)



Obviously, similar results hold for spacetimes which have a spacelike Killing vector; if the Killing vector is taken to be, say,  $\xi^j = \delta_1^j$  in a coordinate system, we can obtain a similar expression starting from the identity  $E_1^1 = 0$ . When the spacetime has more than one Killing vector, we obtain, in general, more than one choice of  $R^j$  which is again an example of the non-uniqueness mentioned earlier.

## 2.6 Boundary terms for the action principle

Part of the interest in the Lanczos-Lovelock models arises from the fact that they admit a standard initial value problem, because the equations of motion contain only up to second derivatives of the metric. At the level of the action, this means that one can hope to derive the equations of motion from a well defined variational principle by fixing the metric on given boundaries, without having to fix also the canonical momenta. For this to happen, one must be able to find suitable boundary terms which depend on the intrinsic geometry of the boundary, the variation of which can cancel the variation of normal derivatives of the metric on the boundary. Such a prescription is needed even for EH action, and the appropriate term is the well known York-Gibbons-Hawking term, given by the integral of the trace of the extrinsic curvature,  $K = h^{\mu\nu} K_{\mu\nu}$ , on the boundary [22]. The question arises as to whether one can construct a similar term for a generic Lanczos-Lovelock action. The answer is yes, and the boundary term itself is most easily obtained in the Hamiltonian formalism, which gives the canonical momenta the trace of which is proportional to the boundary term. For want of a better name, we shall call the Lanczos-Lovelock boundary terms as GYGH (Generalized York-Gibbons-Hawking). The first such analysis, to the best of our knowledge, seems to have been provided by Teitelboim and Zanelli [23].

Although we shall not discuss the details of the Hamiltonian formalism for Lanczos-Lovelock, it is relevant to mention one aspect of the same. For Lanczos-Lovelock action, the canonical momenta  $\pi_{\mu\nu}$  can indeed be found as a function of extrinsic and intrinsic curvature of the boundaries, but the corresponding expression cannot be inverted in closed form to find  $K_{\mu\nu}$  in terms of  $\pi_{\mu\nu}$  (i.e., one cannot express “velocities” in terms of “conjugate momentum”). More importantly,  $K_{\mu\nu}$  would generically be multiple valued functions of  $\pi_{\mu\nu}$  for arbitrary Lanczos-Lovelock couplings (the  $c_m$ ’s). Hence the Hamiltonian evolution is not unique, and the generators of diffeomorphisms cannot be expressed in terms of canonical variables. We refer the interested reader to Ref.[23] and Ref. [24] for further details on the Hamiltonian formulation of Lanczos-Lovelock models. Since we will be working with the Lagrangian formalism all along in this review, the expression for boundary terms in terms of  $K_{\mu\nu}$  is sufficient for the specific topics we wish to cover in later parts of this review. Hence we shall focus on the same.

In what follows, we give the boundary term for a timelike boundary. The corresponding expression for a spacelike boundary, which is the case dealt with in Ref.[23], is similar except for a few (non-trivial) sign changes which one must be careful about. The timelike case turns out to be more relevant for study of quantum black holes since one usually describes the near-horizon physics by replacing the horizon with a *timelike* surface – the stretched horizon – in the spirit of the membrane paradigm. In fact, the boundary terms given in this section will be used later in Secs. 3.2, 3.3 where we will obtain some important conceptual insights into horizon entropy essentially by studying the near-horizon structure of these boundary terms.

For the Lanczos-Lovelock models, the final form of the relevant boundary term can be written as [25]

$$\begin{aligned}
A_{\text{sur(m)}} &= \int_{\partial M} d^{(D-1)}x \sqrt{-h} C_m \\
C_m &= 2m \int_0^1 ds \delta_{j_1 j_2 j_3 \dots j_{2m-1}}^{i_1 i_2 i_3 \dots i_{2m-1}} K_{i_1}^{j_1} \left( \frac{1}{2} R_{i_2 i_3}^{j_2 j_3} - s^2 K_{i_2}^{j_2} K_{i_3}^{j_3} \right) \dots \left( \frac{1}{2} R_{i_{2m-2} i_{2m-1}}^{j_{2m-2} j_{2m-1}} - s^2 K_{i_{2m-2}}^{j_{2m-2}} K_{i_{2m-1}}^{j_{2m-1}} \right)
\end{aligned} \tag{55}$$

The above form can be further simplified, and the complete action can be written as

$$S = \sum_m \left\{ \int_M d^D x \sqrt{-g} c_m L_m \pm \underbrace{\frac{2}{D-2m} \int_{\partial M} d^{(D-1)}x \sqrt{-h} h_{\mu\nu} t_{(m)}^{\mu\nu}}_{A_{\text{sur(m)}}} \right\} \tag{56}$$

where

$$\begin{aligned}
8\pi t_{(m)\mu}^\nu &= \frac{c_m m!}{2^{m+1}} \sum_{s=0}^{m-1} \tilde{C}_s \tilde{\pi}_{(s)\mu}^\nu \\
\tilde{\pi}_{(s)\mu}^\nu &= \delta_{[\mu\mu_1 \dots \mu_{2m-1}]}^{[\nu\nu_1 \dots \nu_{2m-1}]} \mathcal{R}_{\nu_1 \nu_2}^{\mu_1 \mu_2} \dots \mathcal{R}_{\nu_{2s-1} \nu_{2s}}^{\mu_{2s-1} \mu_{2s}} K_{\nu_{2s+1}}^{\mu_{2s+1}} \dots K_{\nu_{2m-1}}^{\mu_{2m-1}}
\end{aligned} \tag{57}$$

and the coefficients  $\tilde{C}_s$  are given by

$$\tilde{C}_s = \frac{4^{m-s}}{s! (2m-2s-1)!!} \tag{58}$$

As the notation already anticipates, the object  $t_{(m)\mu}^\nu$  defined above can be shown to be the surface stress tensor in Lanczos-Lovelock theory; that is,

$$t_{\mu\nu}^{(m)} = \frac{2}{\sqrt{|h|}} \frac{\delta A_{\text{sur(m)}}}{\delta h_{\mu\nu}} \tag{59}$$

In the above expressions,  $\mathcal{R}_{\mu\rho\sigma\nu}$  denotes the *projection* of full spacetime curvature tensor on the  $(D-1)$  hypersurface  $\partial M$ . It is more convenient to express the stress tensor in terms of the *intrinsic* curvature  $\hat{R}_{\mu\rho\sigma\nu}$  of the hypersurface (i.e., the one defined using the induced metric  $h_{\mu\nu}$ ), through the Gauss-Codazzi relation

$$\hat{R}_{\mu\rho\sigma\nu} = \mathcal{R}_{\mu\rho\sigma\nu} - K_{\mu\sigma} K_{\rho\nu} + K_{\mu\nu} K_{\rho\sigma} \tag{60}$$

After some simple combinatorial manipulations, the surface stress energy tensor can be rewritten as

$$\begin{aligned}
8\pi t_{(m)\mu}^\nu &= \frac{c_m m!}{2^{m+1}} \sum_{s=0}^{m-1} C_s \pi_{(s)\mu}^\nu \\
\pi_{(s)\mu}^\nu &= \delta_{[\mu\mu_1 \dots \mu_{2m-1}]}^{[\nu\nu_1 \dots \nu_{2m-1}]} \hat{R}_{\nu_1 \nu_2}^{\mu_1 \mu_2} \dots \hat{R}_{\nu_{2s-1} \nu_{2s}}^{\mu_{2s-1} \mu_{2s}} K_{\nu_{2s+1}}^{\mu_{2s+1}} \dots K_{\nu_{2m-1}}^{\mu_{2m-1}}
\end{aligned} \tag{61}$$

with

$$C_s = \sum_{q=s}^{m-1} \frac{(-2)^{q-s} 4^{m-q} \binom{q}{s}}{q! (2m - 2q - 1)!!} \quad (62)$$

One can check that for  $m = 1$  and  $m = 2$ , the above expression reduces to that of the surface stress tensor in Einstein [26] and Gauss-Bonnet [27] theories respectively.

Although we have quoted the structure of the surface terms, we have not formally shown that these terms indeed make the variational principle well defined. In fact, this is a remarkable feature of the Lanczos-Lovelock actions and the reason why things work out again lies in the topological origin of the Lanczos-Lovelock actions. We refer the reader to [19, 28] for a detailed description of this point. It is however worth giving a sketch of the main argument here.

As we mentioned in the previous section, in the critical dimension  $D = 2m$ , the Lanczos-Lovelock action  $L_m^{(2m)}$  gives Euler character of a manifold  $\mathcal{M}$  *without* a boundary; in presence of a boundary  $\partial\mathcal{M}$ , one needs to add a specific boundary term, say  $-Q_m^{(2m)}$ , to obtain the correct Euler characteristic. Since Euler characteristic can not change under smooth variations of a metric with a fixed boundary, the variation of  $\int_{\mathcal{M}} L_m^{(2m)} - \int_{\partial\mathcal{M}} Q_m^{(2m)}$  with respect to normal derivatives of the metric on the boundary must vanish; in other words, the corresponding variations in  $L_m^{(2m)}$  and  $Q_m^{(2m)}$  must cancel. This *cancellation* is independent of spacetime dimensions, and hence the above result is valid even away from the critical dimension. So the boundary term we need is essentially the term which — when added to the bulk term — will make the sum equal to the Euler characteristic in the critical dimension. From this one can show that the form of the surface term is given by the Chern-Simons form defined on  $\partial\mathcal{M}$ . (The above points are easily understood in Einstein gravity, where the standard York-Gibbons-Hawking term is precisely what is needed to make the full action Euler characteristic in  $D = 2$ , which is the critical dimension for Einstein-Hilbert term corresponding to  $m = 1$ .) This leads to the boundary terms quoted above.

We stress a few important characteristics of the above surface term, some of which will be relevant from the point of view of topics to be discussed in Sec. 3.3.

1. The surface term for a given Lanczos-Lovelock term of order  $m$  has an apparent pole at the critical dimension  $D = 2m$ . There is, however, no real pole since  $t_{(m)\mu}^\nu$  itself has a zero at  $D = 2m$  which precisely cancels the pole. This will also become evident when we give explicit form for  $t_{(m)\mu}^\nu$  near a static horizon in Sec. 3.3. In fact, what happens at  $D = 2m$  (which is the critical dimension) is straightforward to understand from the fact that, in  $D = 2m$  the corresponding Lanczos-Lovelock term is the Euler density of the manifold, and the above surface term must be included if the manifold has a boundary.
2. For a given  $m$ , it is easy to see that the boundary term is a polynomial containing *odd* powers of  $K_{\mu\nu}$  of degree 1 to  $(2m - 1)$ . For e.g., for the Gauss-Bonnet term ( $m = 2$ ), the boundary term will be a linear combination of terms linear and cubic in  $K_{\mu\nu}$ .
3. The notation for the boundary tensor anticipates the fact that the *jump* in  $t_{\mu\nu}$  across a shell, say, can be related to stress-energy tensor of the material comprising the shell, to yield the (generalization of) Israel junction conditions for Lanczos-Lovelock models.

This fact is useful in the study of collapse and black hole formation in Lanczos-Lovelock models. Further, the same junction conditions can be used to set-up the membrane paradigm for black hole solutions in Lanczos-Lovelock theory, in the sense that one can replace the details of region hidden by the horizon by endowing the stretched horizon with the above stress tensor. We expect the membrane paradigm to work in this case just as it does in Einstein theory, due to the initial value problem being well defined in Lanczos-Lovelock models. We will discuss this further in section 3.3.

## 2.7 Holographic structure of Lanczos-Lovelock action

In a classical theory, action functional is just a device to obtain equations of motion and to encode the symmetries of the system. The actual value of the action functional is of no importance. In contrast, action functional plays a more crucial role in quantum theory: in the semiclassical limit, it gives the phase of the wave functional and in the full description it appears in the path integral. So the off-shell structure of the action functional is important in quantum theory. In Lanczos-Lovelock models we have already seen that the action has a peculiar structure and requires a special treatment if it has to lead to a well-defined variational principle. By and large, this requires careful handling of the surface term in the action functional.

The existence of the surface term brings in another aspect in the Lanczos-Lovelock models viz. that the bulk and surface terms of the action are closely related to each other and duplicates the information to certain extent. This is reminiscent of the notion of holography, which represents — in the most general sense — a correspondence between bulk and boundary degrees of freedom. For lack of better terminology we will call the *correspondence between bulk and boundary terms in the action* (rather than in the degrees of freedom) as "holography of the action functional". (In this review, we shall use the term 'holography' as explained above and it should not be confused with the usage of this term in e.g., string theory.) In this section, we shall explore this feature and show that Lanczos-Lovelock actions possess a holographic structure which eventually gets reflected in the horizon thermodynamics in these models.

The basic idea we wish to build upon was already expressed, symbolically, in the decomposition of the Lagrangian in Eq. (27). We will first describe a toy example of higher derivative action in ordinary mechanics to demonstrate the fact that such an action must have a very special structure if it has to yield equations of motion which are still second order in time derivatives [29]. It is then not a surprise that the generalization to gravitational actions picks out the Lanczos-Lovelock action uniquely as the one exhibiting holography.

### 2.7.1 Higher derivative actions yielding second order equations of motion

Consider a dynamical variable  $q(t)$  in classical mechanics described by a Lagrangian  $L_q(q, \dot{q})$ . Varying the action obtained from integrating this Lagrangian in the interval  $(t_1, t_2)$  while keeping  $q$  fixed at the endpoints, we obtain the Euler Lagrange equations for the system  $(\partial L_q / \partial q) = dp/dt$ , where  $p(q, \dot{q}) \equiv (\partial L_q / \partial \dot{q})$ . (The subscript  $q$  on  $L_q$  denotes the variable that is kept fixed at the end points.) The Lagrangian depends only up to first derivatives of the dynamical variable and the equations of motion are — in general — second degree in the time derivative.

Interestingly, there exists a wide class of Lagrangians  $L(\ddot{q}, \dot{q}, q)$  which depend on  $\ddot{q}$  as well but still lead to equations of motion which are only second order in time. We will now analyse this class which will generalize to the holographic actions in field theory.

To do this, let consider the following question: We want to modify the Lagrangian  $L_q$  such that the same equations of motion are obtained when — instead of fixing  $q$  at the end points — we fix some other (given) function  $C(q, \dot{q})$  at the end points. This is easily done by modifying the Lagrangian by adding a term  $-df(q, \dot{q})/dt$  which depends on  $\dot{q}$  as well. (The minus sign is for future convenience.) The new Lagrangian is:

$$L_C(q, \dot{q}, \ddot{q}) = L_q(q, \dot{q}) - \frac{df(q, \dot{q})}{dt} \quad (63)$$

We want  $L_C$  to lead to the same equations of motion as  $L_q$ , when some given function  $C(q, \dot{q})$  is held fixed at the end points. We assume  $L_q$  and  $C$  are given and that we need to find  $f$ . The standard variation now gives

$$\delta A_C = \int_{t_1}^{t_2} dt \left[ \left( \frac{\partial L}{\partial q} \right) - \frac{dp}{dt} \right] \delta q - \int_{t_1}^{t_2} dt \frac{d}{dt} [\delta f - p \delta q] \quad (64)$$

At the end-points, we invert the relation  $C = C(q, \dot{q})$  to determine  $\dot{q} = \dot{q}(q, C)$  and express  $p(q, \dot{q})$  in terms of  $(q, C)$  obtaining the function  $p = p(q, C)$ . We then treat  $f$  as a function of  $q$  and  $C$ , so that the variation of the action becomes:

$$\begin{aligned} \delta A_C &= \int_{t_1}^{t_2} dt \left[ \left( \frac{\partial L}{\partial q} \right) - \frac{dp}{dt} \right] \delta q + \left[ p(q, C) - \left( \frac{\partial f}{\partial q} \right)_C \right] \delta q \Big|_{t_1}^{t_2} - \left( \frac{\partial f}{\partial C} \right)_q \delta C \Big|_{t_1}^{t_2} \\ &= \int_{t_1}^{t_2} dt \left[ \left( \frac{\partial L}{\partial q} \right) - \frac{dp}{dt} \right] \delta q + \left[ p(q, C) - \left( \frac{\partial f}{\partial q} \right)_C \right] \delta q \Big|_{t_1}^{t_2} \end{aligned} \quad (65)$$

since  $\delta C = 0$  at the end points by assumption. To obtain the same Euler-Lagrange equations, the second term in Eq. (65) should vanish for all  $\delta q$ . This fixes the form of  $f$  to be:

$$f(q, C) = \int p(q, C) dq + F(C) \quad (66)$$

where the integration is with constant  $C$  and  $F$  is an arbitrary function.

Thus, given a Lagrangian  $L_q(q, \dot{q})$  which leads to certain equations of motion when  $q$  is held fixed, one can construct a *family* of Lagrangians  $L_C(q, \dot{q}, \ddot{q})$  that will lead to the *same* equations of motion when an arbitrary function  $C(q, \dot{q})$  is kept fixed at the end points. *This family is remarkable in the sense that though  $L_C$  will now also involve  $\ddot{q}$ , the equations of motion are still of second order in  $q$  because of the structure of the Lagrangian.* (The results obtained above can also be interpreted in terms of canonical transformations etc. which we purposely avoid since we want to stay within the Lagrangian framework). So, even though a *general* Lagrangian which depends on  $\ddot{q}$  will lead to equations of higher order, there are *special* Lagrangians which will not. The analysis extends directly to a multicomponent field  $q_A(x^a)$  in a spacetime with coordinates  $x^a$  where  $A$  collectively denotes the tensor indices [29].

When one considers the action just as a tool to obtain the field equations, the above procedure can be used with any  $C$ . But once the dynamical variables in the theory have been

identified, there are two natural boundary conditions which can be imposed on the system. The first one holds  $q$  fixed at the boundary and the second one keeps the canonical momenta  $p$  fixed at the boundary. In the second case,  $C = p$  and Eq.(66) gives  $f = qp = q(\partial L/\partial \dot{q})$  leading to the Lagrangian:

$$L_p = L_q - \frac{d}{dt} \left( q \frac{\partial L_q}{\partial \dot{q}} \right) . \quad (67)$$

This Lagrangian  $L_p$  will lead to the same equations of motion when  $p(q, \dot{q}) = (\partial L/\partial \dot{q})$  is held fixed at the end points. In field theory, with multicomponent field  $q_A(x^a)$  with canonical momenta  $\pi_A^i = [\partial L/\partial(\partial_i q^A)]$ , when we fix  $C_A^a = \pi_A^a$  at the boundary, the new Lagrangian we need to use is:

$$L_p(\partial^2 q_A, \partial q_A, q_A) = L_q(q_A, \partial q_A) - \partial_i \left[ q_A \left( \frac{\partial L_q}{\partial(\partial_i q^A)} \right) \right] \equiv L_{\text{bulk}} + L_{\text{sur}} \quad (68)$$

We will now summarise [29] the implications of the above results:

1. These results have a natural interpretation in quantum theory: A path integral defined with  $L_p$  will give the transition amplitude in momentum space  $G(p_2, t_2; p_1, t_1)$ , just as a path integral with  $L_q$  leads to the transition amplitude in coordinate space  $K(q_2, t_2; q_1, t_1)$ .
2. When the momenta are held fixed, the surface term has, what we shall call, the “ $d(qp)$ ” structure. It is obvious from this relation that the surface and bulk terms in the action are closely related and the form of the surface term will put some constraints on  $L_{\text{bulk}} = L_q$ .
3. It turns out that the Einstein-Hilbert action (and more generally the Lanczos-Lovelock action) has precisely this form, which is evident when the action is written in the  $\Gamma\Gamma$  form. *This is the key to the holography in Lanczos-Lovelock action functionals which we now describe.*

### 2.7.2 Holography of gravitational action functionals

Let us first discuss the Einstein-Hilbert action, rewriting the well known  $\Gamma\Gamma$  form of the action in a manner which generalizes to the Lanczos-Lovelock actions. The Einstein-Hilbert action can be written using a tensor  $Q_{abcd} \equiv (1/2)(g_{ac}g_{bd} - g_{ad}g_{bc})$  as:

$$\begin{aligned} \sqrt{-g}L_{\text{EH}} &= Q^{abcd}R_{abcd} \\ &= 2\partial_c [\sqrt{-g}Q_a{}^{bcd}\Gamma_{bd}^a] + 2\sqrt{-g}Q_a{}^{bcd}\Gamma_{dk}^a\Gamma_{bc}^k \\ &\equiv L_{\text{sur}} + L_{\text{bulk}} \end{aligned} \quad (69)$$

with

$$L_{\text{bulk}} = 2\sqrt{-g}Q_a{}^{bcd}\Gamma_{dk}^a\Gamma_{bc}^k ; \quad L_{\text{sur}} = 2\partial_c [\sqrt{-g}Q_a{}^{bcd}\Gamma_{bd}^a] \equiv \partial_c [\sqrt{-g}V^c] \quad (70)$$

where the last equality defines the  $D$ -component non-tensor  $V^c$ . As is well known, one can obtain Einstein's equations varying only  $L_{\text{bulk}}$  keeping  $g_{ab}$  fixed at the boundary.

The non-trivial aspect relevant for holography is a specific relation [30] between  $L_{\text{bulk}}$  and  $L_{\text{sur}}$  allowing  $L_{\text{sur}}$  to be determined completely by  $L_{\text{bulk}}$ . Using  $g_{ab}$  and  $\partial_c g_{ab}$  as independent variables in  $L_{\text{bulk}}$  one can prove that:

$$L_{\text{sur}} = - \frac{1}{[(D/2) - 1]} \underbrace{\partial_i \left( g_{ab} \frac{\partial L_{\text{bulk}}}{\partial (\partial_i g_{ab})} \right)}_{\text{"d}(qp)"} \quad (71)$$

The connection with the discussion in the last subsection is obvious: the “ $d(qp)$ ” structure of  $L_{\text{sur}}$  suggests that  $L_{\text{EH}}$  is obtained from  $L_{\text{bulk}}$  by a transformation from coordinate space to momentum space. The action based on  $L_{\text{bulk}}$  is the co-ordinate space action and leads to a well-defined variational principle if the metric is fixed at the boundary. On the other hand, the usual Einstein-Hilbert action ( $R\sqrt{-g} = L_{\text{bulk}} + L_{\text{sur}}$ ) is actually a *momentum-space* action and leads to well-defined variational principle if the momenta are fixed at the boundary. This is very transparent if we use  $f^{ab} \equiv \sqrt{-g}g^{ab}$  as the dynamical variables with the momenta being

$$N_{jk}^i \equiv \frac{\partial(\sqrt{-g}L_{\text{bulk}})}{\partial(\partial_i f^{jk})} = -[\Gamma_{jk}^i - \frac{1}{2}(\delta_j^i \Gamma_{ka}^a + \delta_k^i \Gamma_{ja}^a)]. \quad (72)$$

The equations of motion will follow from  $\delta A_{EH} = 0$  if we fix the momenta  $N_{jk}^i$  on the boundary. In fact, if we decide *not* to add any surface term to Einstein-Hilbert action, then we can still obtain [5] the field equations if we demand:

$$\delta A_{EH} = - \int_{\partial V} d^3x \sqrt{h} n_i g^{jk} \delta N_{jk}^i \quad (73)$$

instead of the usual  $\delta A_{EH} = 0$ .

Similar results continue to hold for Lanczos-Lovelock actions, which we will now describe. Taking a cue from Einstein’s theory, consider a Lagrangian of the form:

$$\sqrt{-g}L = 2\partial_c [\sqrt{-g}Q_a^{bcd}\Gamma_{bd}^a] + 2\sqrt{-g}Q_a^{bcd}\Gamma_{dk}^a\Gamma_{bc}^k \equiv L_{\text{sur}} + L_{\text{bulk}} \quad (74)$$

where  $Q^{abcd}$  is a tensor having the symmetries of curvature tensor. By straight forward algebra, one can prove [29] the following identity:

$$\sqrt{-g}L = 2\partial_c [\sqrt{-g}Q_a^{bcd}\Gamma_{bd}^a] + 2\sqrt{-g}Q_a^{bcd}\Gamma_{dj}^a\Gamma_{bc}^j = \sqrt{-g}Q_a^{bcd}R_{bcd}^a + 2\sqrt{-g}\Gamma_{bd}^a\nabla_c Q_a^{bcd} \quad (75)$$

The general covariance of  $L$  requires the condition  $\nabla_c Q_a^{bcd} = 0$ . Then we get:

$$\sqrt{-g}L = \sqrt{-g}Q_a^{bcd}R_{bcd}^a; \quad \nabla_c Q_a^{bcd} = 0 \quad (76)$$

We have, therefore, once again arrived at the Lanczos-Lovelock models in a very different context, both mathematically and conceptually. The final step of establishing *holography* of such actions requires some additional care particularly since, unlike in the case of Einstein theory,  $L_{\text{bulk}}$  will in general depend on second derivatives of the metric. For Lanczos-Lovelock models, it is possible to show that [29] the following generalization of holography exists:

$$[(D/2) - m]L_{\text{sur}} = -\partial_i \left\{ \left[ g_{ab} \frac{\delta}{\delta(\partial_i g_{ab})} + \partial_j g_{ab} \frac{\partial}{\partial(\partial_i \partial_j g_{ab})} \right] L_{\text{bulk}} \right\} \quad (77)$$

where the *Eulerian derivative*  $\delta$  is defined as

$$\frac{\delta K[\phi, \partial_i \phi, \dots]}{\delta \phi} = \frac{\partial K[\phi, \partial_i \phi, \dots]}{\partial \phi} - \partial_a \left[ \frac{\partial K[\phi, \partial_i \phi, \dots]}{\partial (\partial_a \phi)} \right] + \dots \quad (78)$$

Note that the second term on right hand side of Eq. (77) vanishes for  $m = 1$  Lanczos-Lovelock term (which is the Einstein theory). The proof of above relation, as well as some other forms of holography for the Lanczos-Lovelock class of actions (which arise by choosing the Christoffel symbols in place of first derivatives of the metric as independent variables), can be found in Ref. [29].

## 2.8 Black hole and cosmological solutions of Lanczos-Lovelock models

There is a vast literature on known solutions for gravitational theories based on the Lanczos-Lovelock Lagrangians, and even a brief discussion of these would form a topic for a new review. We shall therefore use this section just to redirect the interested reader to relevant literature and reviews where more details about specific solutions, and more importantly, additional references, can be found.

As far as black hole solutions of the theory are considered, one of the very first solution to have appeared (to the best of our knowledge), is the so called Boulware-Deser solution [31] which is essentially a spherically symmetric black hole solution of 5 dimensional Einstein-Gauss-Bonnet (EGB) action. The metric for this solution is given by

$$\begin{aligned} ds^2 &= -f(r)dt^2 + \frac{1}{f(r)}dr^2 + r^2 d\Omega_3^2 \\ f(r) &= 1 + \frac{r^2}{4\alpha} \left[ 1 \pm \sqrt{1 + \frac{16\alpha m}{r^4} + \frac{4\alpha\Lambda}{3}} \right] \end{aligned} \quad (79)$$

where  $d\Omega_3^2$  is the metric on a unit 3-sphere,  $\Lambda$  is the cosmological constant, and  $\alpha$  is the Gauss-Bonnet coupling constant normalised so that the total lagrangian takes the form:  $L \propto -2\Lambda + R + \alpha (R^2 - 4R_{ab}R^{ab} + R_{abcd}R^{abcd})$ . The parameter  $m$  is an integration constant which is proportional to the “mass” of the solution; its exact form with all the constants, can be found, for example, in [31]. The  $\pm$  is a consequence of the fact that field equations are quadratic in  $f(r)$ . There exists an extensive literature on properties of this solution, as well as its modification upon addition of the electric charge  $Q$ .

There also exists considerable literature on study of Birkhoff’s theorem in EGB theory, the most relevant of which are listed in [32]. Although the analysis leading to the Boulware-Deser solution presented in [31] was motivated by appearance of EGB term in low energy expansion of super-symmetric string theory, it was followed by extensive work on finding such solutions for generic Lanczos-Lovelock models treated as classical models of higher dimensional gravity in their own right. A nice review of such solutions, and their various asymptotic properties (which are of some interest also for the AdS-CFT correspondence), can be found in Refs. [33] and [34] (see also [35] for work generalizing known solutions in Einstein theory to Lanczos-Lovelock). A systematic study of *cosmological solutions* in the Lanczos-Lovelock theory can be found in [36], which deals in particular with maximally



symmetric space-times, Robertson-Walker universes, and product manifolds of symmetric subspaces.

### 3 Lanczos-Lovelock models: Horizon thermodynamics

One fascinating aspect of general relativity, which is still not completely understood conceptually, is the association of thermodynamic variables with horizons. Originally, these results were obtained in the context of black hole physics [37, 38] but it is now clear that all null surfaces possess these attributes in an observer-dependent manner [39]. It is important to understand whether this is a special feature of Einstein's theory or whether all Lanczos-Lovelock models share this intriguing connection. It turns out, as we shall describe in this and ensuing sections, that the latter is the case.

As we said before, one key utility of the Lanczos-Lovelock models is in their power to discriminate between widely differing conceptual explanations for known results in Einstein gravity. So the study of horizon thermodynamics in Lanczos-Lovelock models will allow us to distinguish those aspects which transcend Einstein's theory and will lead to a more promising and general route to understanding these aspects. In this spirit, there has been an extensive study of thermodynamics of black hole solutions in particular and horizons in general in the Lanczos-Lovelock theory which we now turn to. Since it is impossible to do justice to all the literature in this review, we will be selective in our discussion. A good introduction to this topic can be found in the (possibly the very first) paper on the topic by Myers and Simon [40].

#### 3.1 Horizon entropy in Lanczos-Lovelock theory

The results from the study of quantum field theory in a given background spacetime does not depend on the action governing the gravitational dynamics of the background. In particular, results [39] such as the Davies-Unruh effect etc continue to hold in a given metric, independent of the field equations of the theory. In fact, the notion of horizon temperature only involves the near horizon geometric properties of the spacetime metric; therefore, the notion of temperature for a horizon remains the same independent of the theory of gravity.

The same is not true for the horizon entropy, though. Algebraically, this is easy to understand in the context of a simple black hole solution in Lanczos-Lovelock theory characterized by a single parameter corresponding to mass  $M$  of the black hole (identified in a suitable way, for example, by looking at the asymptotics of the solution). Then, the entropy can be identified by integrating  $dM/T(M)$ , for which one would need an explicit form for  $T(M)$ . This depends on the equations of motion of the theory.

There exists a general way of identifying horizon entropy in higher derivative theories, given by Wald [41] and Visser [42, 43], which makes the dependence of entropy on gravitational action very evident. In this approach, horizon entropy is identified as a surface integral over the horizon of a  $(D-2)$ -form corresponding to the Noether charge arising from diffeomorphism invariance of the theory. The basic idea is to find the Noether current  $\mathbf{J}$  of diffeomorphism invariance (which is a closed  $(D-1)$  form), and use the associated Noether potential  $\mathbf{Q}$  defined by  $\mathbf{J} = d\mathbf{Q}$  to define black hole entropy. By considering variations (not necessarily stationary) of a stationary background spacetime possessing a Killing horizon, Wald argued

that the first law of black hole dynamics is consistent with the identification of the following expression as the horizon entropy:

$$S_{\text{Noether}} = \beta \int d^{D-1} \Sigma_a J^a = \frac{1}{2} \beta \int d^{D-2} \Sigma_{ab} J^{ab} \quad (80)$$

where  $\kappa$  is the surface gravity,  $\beta^{-1} = 2\pi/\kappa$  is the temperature of the horizon and  $J^a, J^{ab}$  are the Noether current and the Noether potential. In the final expression in Eq. (80) the integral is over any  $(D-2)$  dimensional spacelike cross-section of the Killing horizon on which the norm of  $\xi^a$  vanishes. For example, in the case of Einstein gravity, Eq. (41) reduces to

$$J^{ab} = \frac{1}{16\pi} (\nabla^a \xi^b - \nabla^b \xi^a) \quad (81)$$

where  $\xi^a$  is the Killing vector generating the horizon. Writing the area element as  $d\Sigma_{ab} = (l_a \xi_b - l_b \xi_a) \sqrt{\sigma} d^{D-2}x$  where  $l_a$  is an auxiliary vector field satisfying  $l_a \xi^a = -1$ , the integral in Eq. (80) reduces to

$$S_{\text{Noether}} = -\beta \frac{1}{8\pi} \int d^{D-2}x \sqrt{\sigma} (l_a \xi_b) \nabla^b \xi^a = \beta \frac{1}{8\pi} \kappa \int d^{D-2}x \sqrt{\sigma} = \frac{1}{4} A_H \quad (82)$$

where  $A_H$  is the horizon area. (We have used the relations  $\xi^a \nabla_a \xi^b = \kappa \xi^b$  and  $\beta \kappa = 2\pi$ .) This, of course, agrees with the Bekenstein-Hawking entropy.

In what follows, we first give a quick derivation (in component notation) of the arguments of Wald as to why the definition of entropy based on Noether potential makes sense for stationary Killing horizons. It turns out that the Noether current  $J^a$  and its associated potential  $J^{ab}$  (or more accurately, their *off-shell* versions) characterize most of the results related to thermodynamic aspects of gravity, and provide the most natural substitute for the equation of motion tensor  $E_{ab}$  in attempts to interpret gravitational dynamics as an emergent phenomenon. So this description is of importance in our later discussions on emergent paradigm.

Let us now sketch the argument leading to an expression for black hole entropy in a general gravitational theory with the Lagrangian  $L[g_{ab}, R_{abcd}]$ . The main input needed is the definition of covariant version of Hamiltonian equations of motion, which can be given in terms of a symplectic  $(D-1)$  form  $\omega^a$ , which is defined, in component notation, as

$$\omega^a \equiv \delta(\sqrt{-g}) \delta_{\xi} v^a - \delta_{\xi}(\sqrt{-g}) \delta v^a \quad (83)$$

where  $\delta_{\xi}$  is the Lie variation with respect to a vector field  $\xi$ , and  $\delta_{\xi} v^a$  is defined from the variation of the Lagrangian under a diffeomorphism:

$$\delta(L\sqrt{-g}) = \sqrt{-g} E_{ab} \delta g^{ab} + \sqrt{-g} \nabla_a (\delta_{\xi} v^a) \quad (84)$$

In terms of the above symplectic form, covariant Hamilton's equations take the form (note that  $\omega^a$  is a tensor density of weight unity)

$$\delta H[\xi] = \int_c d\Sigma_a \omega^a / \sqrt{-g} \quad (85)$$

where  $\mathcal{C}$  is a Cauchy surface.

A rigorous justification for this can be found in [44], which studies the covariant phase space formulation for field theories. We here take a digression to simply motivate the above expression in analogy with point particle mechanics. In this case, the symplectic form is  $\Theta = \mathbf{d}p \wedge \mathbf{d}q$ , in terms of which Hamilton's equations can be written as

$$\begin{aligned} \mathbf{d}H &= \Theta(\dots, \partial_t) \\ &= \dot{q}(q, p) \mathbf{d}p - \dot{p}(q, p) \mathbf{d}q \end{aligned} \quad (86)$$

The Hamiltonian  $H$  here generates time translations (hence the vector  $\partial_t$  in the 2nd slot of  $\Theta$ ), while the “d”'s are arbitrary variations whose coefficients characterize the  $(q, p)$  dependence of the Hamiltonian. One may re-write the above expression as follows.

$$\begin{aligned} \mathbf{d}H &= \dot{q} \mathbf{d}p - \dot{p} \mathbf{d}q \\ &= \mathbf{d}(p\dot{q}) - p\mathbf{d}\dot{q} - \dot{p}\mathbf{d}q \\ &= \mathbf{d}(p\mathbf{d}_t q) - \mathbf{d}_t(p\mathbf{d}q) \end{aligned} \quad (87)$$

Now, the surface term in point particle action is  $v(\mathbf{d}q) = p\mathbf{d}q$ . Hence we can write

$$\mathbf{d}H = \mathbf{d}[v(\dot{q})] - \mathbf{d}_t[v(\mathbf{d}q)] \quad (88)$$

which has the same structure as Eq. (85) with  $\omega^a$  given in Eq. (83). (Similar conclusion would be reached if the surface term is  $q\mathbf{d}p$ , with the roles of  $q$  and  $p$  interchanged.)

Returning now to the main discussion, we see that  $\omega^a$  can be written as:

$$\begin{aligned} \omega^a &= \delta(\sqrt{-g} \delta_{\xi} v^a) - \delta_{\xi}(\sqrt{-g} \delta v^a) \\ &= \sqrt{-g} \left[ \delta(\delta_{\xi} v^a) - \mathcal{L}_{\xi}(\delta v^a) \right] + (\delta_{\xi} v^a)(\delta\sqrt{-g}) - (\delta v^a)(\delta_{\xi}\sqrt{-g}) \end{aligned} \quad (89)$$

This allows us to relate  $\omega^a$  to the Noether current  $J^a$ , obtained from diffeomorphism invariance of the Lagrangian, as given by Eq. (40) which is conserved off-shell. If we omit from it the term involving  $E_{ab}$  we get the expression:

$$J^a[\xi] = \delta_{\xi} v^a - L\xi^a \quad (90)$$

which is conserved *on-shell*, i.e., when the equations of motion are satisfied:

$$\begin{aligned} \nabla_a J^a &= \nabla_a(\delta_{\xi} v^a) - \nabla_a(L\xi^a) \\ &= \frac{1}{\sqrt{-g}} \mathcal{L}_{\xi}(L\sqrt{-g}) - E_{ab}\delta g^{ab} - \nabla_a(L\xi^a) \\ &= -E_{ab}\delta g^{ab} \end{aligned} \quad (91)$$

since  $\mathcal{L}_{\xi}(L\sqrt{-g}) = \sqrt{-g}\nabla_a(L\xi^a)$ . Therefore,  $\nabla_a J^a = 0$  when  $E_{ab} = 0$ . We will use the definition of  $J^a$  in the expression for  $\omega^a$  and also utilize the facts: (i)  $\delta_{\xi} v^a = J^a[\xi] + L\xi^a$ , (ii)  $\delta(\sqrt{-g} L) = \sqrt{-g} \nabla_a \delta v^a$  on-shell, (iii)  $\delta_{\xi}\sqrt{-g} = \sqrt{-g}\nabla_a \xi^a$ , and (iv)  $\delta$  represents arbitrary

field variations, and therefore do not act on  $\xi$ . Straightforward algebra then leads to the result:

$$\omega^a = \delta (J^a[\xi]\sqrt{-g}) + 2\sqrt{-g} \nabla_b (\xi^{[a}\delta v^{b]}) \quad (92)$$

which can be substituted in Eq. (85) to obtain

$$\delta H[\xi] = \delta \int_{\mathcal{C}} d\Sigma_a J^a + \int_{\mathcal{C}} d\Sigma_a 2\nabla_b (\xi^{[a}\delta v^{b]}) \quad (93)$$

(The steps leading to the above relation are best understood by doing a 3 + 1 decomposition w.r.t. normal  $n_a = \pm N \partial_a n$  to the hypersurface  $\mathcal{C}$ , with  $\pm$  corresponding to spacelike/timelike surface respectively, and using  $\sqrt{-g} = N\sqrt{|h|}$  with  $h_{\mu\nu}$  being the induced metric on  $\mathcal{C}$ .) From the above, one can already note that the flux of Noether current is almost the same as the Hamiltonian, except for a surface term. Further, if the variations  $\delta g_{ab}$  satisfy the linearised equations of motion, then the perturbed current is also conserved, and one therefore has  $\delta[J^a\sqrt{-g}] = \partial_b[\delta(J^{ab}\sqrt{-g})]$ . Repeating the steps leading to the preceding expression, this time in terms of  $J^{ab}$ , leads to

$$\begin{aligned} \delta H[\xi] &= \delta \int_{\mathcal{C}} d\Sigma_a \nabla_b J^{ab} + \int_{\mathcal{S}_\infty} d\Sigma_{ab} \xi^{[a}\delta v^{b]} \\ &= -\frac{1}{2} \delta \int_{\mathcal{S}_{\text{bif}}} d\Sigma_{ab} J^{ab} + \frac{1}{2} \delta \int_{\mathcal{S}_\infty} d\Sigma_{ab} J^{ab} + \int_{\mathcal{S}_\infty} d\Sigma_{ab} \xi^{[a}\delta v^{b]} \end{aligned} \quad (94)$$

where  $\mathcal{S}_\infty$  is the  $(D-2)$  dimensional boundary of  $\mathcal{C}$  at asymptotic infinity. The inner boundary of  $\mathcal{C}$  is taken to be the bifurcation surface  $\mathcal{B}$ , where  $\xi^a = 0$  (hence this surface does not contribute as far as the  $\xi^{[a}\delta v^{b]}$  term is concerned). It is clear from the second expression above that, for a meaningful Hamiltonian  $H$  to exist, the last integral must be expressible as a variation of a  $(D-1)$  form, say  $\mathbf{B}$ . Assuming this is so, then, as argued in [41], the combined contribution to  $H$  of this term and the second term involving  $J^{ab}$  above, at asymptotic infinity  $\mathcal{S}_\infty$ , gives the numerical value of the ADM Hamiltonian when the vector field  $\xi$  is chosen appropriately to correspond to asymptotic time translation and rotation(s). Now, when  $\xi$  is a Killing field, from its very construction,  $\delta H[\xi] = 0$ , although  $H[\xi]$  itself need not vanish. Using all these, one may identify the  $\mathcal{S}_\infty$  contributions above as the “work terms”  $\delta\mathcal{E} - \Omega_H \delta\mathcal{J}$ , with  $\mathcal{E}$  and  $\mathcal{J}$  being the energy and angular momentum respectively (see [41] for a rigorous justification of this). This in turn allows us to identify the black hole entropy  $S_{\text{Noether}}$  from the  $\mathcal{S}_{\text{bif}}$  contribution via the relation  $(\kappa/2\pi)\delta S = \delta\mathcal{E} - \Omega_H \delta\mathcal{J}$ . Symbolically,

$$0 = \underbrace{-\frac{1}{2} \delta \int_{\mathcal{S}_{\text{bif}}} d\Sigma_{ab} J^{ab}}_{-(\kappa/2\pi)\delta S} + \underbrace{\frac{1}{2} \delta \int_{\mathcal{S}_\infty} d\Sigma_{ab} J^{ab} + \int_{\mathcal{S}_\infty} d\Sigma_{ab} \xi^{[a}\delta v^{b]}}_{\delta\mathcal{E} - \Omega_H \delta\mathcal{J}} \quad (95)$$

Using all the above arguments, Wald eventually identified the entropy as

$$S_{\text{Noether}} = \frac{1}{2} \beta \int J^{ab} d\Sigma_{ab} \quad (96)$$

The above expression for entropy can be rewritten in a simplified (and appealing) form for stationary horizons in Lanczos-Lovelock models. However, for sake of simplicity, we sketch how this form is arrived at for static horizons. Using Eq. (45) for  $J_{ab}$  in Lanczos-Lovelock models, we can rewrite Eq. (96) as

$$S_{\text{Noether}} = \frac{1}{2}\beta \int (2P^{abcd} \nabla_{[c}\xi_{d]}) (2\xi_{[a}l_{b]}) \sqrt{\sigma} d^{(D-2)}x \quad (97)$$

where  $l_a$  is an auxiliary null vector satisfying  $l^2 = 0$  and  $l \cdot \xi = -1$ , in terms of which  $d\Sigma_{ab} = (2\xi_{[a}l_{b]}) \sqrt{\sigma} d^{(D-2)}x$  because  $\xi$  is the horizon generating Killing vector field. One can now use the near horizon static structure of the metric (discussed in detail in section Sec. 4.1.1; see Eq. (197)) to show that  $\nabla_{[c}\xi_{d]} = (-2\kappa^2 n) \delta_{[c}^n \delta_{d]}^0$  and  $l^a = (1/N^2, 1/N, 0, \dots)$  near the horizon defined by  $n = 0$  whence the lapse behaves as  $N = \kappa n + O(n^3)$ . Using these, it is easily seen that

$$S_{\text{Noether}} = \frac{1}{2}\beta \int (8\kappa P_{0n}^{0n}) \sqrt{\sigma} d^{(D-2)}x \quad (98)$$

From the total anti-symmetry of the alternating determinant tensor, and the fact that  ${}^{(D-2)}R_{CD}^{AB} = R_{CD}^{AB} + O(n^2)$ , it follows that

$$P_{0n}^{0n} = \left(\frac{m}{2}\right) L_{m-1}^{(D-2)} \quad (99)$$

which eventually gives

$$S_{\text{Noether}} = 4\pi m \int L_{m-1}^{(D-2)} \sqrt{\sigma} d^{(D-2)}x \quad (100)$$

Thus the entropy of  $m$ -th order Lanczos-Lovelock theory in  $D$ -dimensions can be related to the *Lagrangian* of the  $(m-1)$ -th order Lanczos-Lovelock theory in  $(D-2)$ -dimensions. The above expression (which holds for stationary horizons as well) provides a description of horizon entropy in Lanczos-Lovelock models purely in terms of intrinsic geometry of the horizon, without any reference to any additional vector fields etc. For a more general (non Lanczos-Lovelock) Lagrangian, this, of course, will not be the case. In particular, the fact that entropy is purely a functional of intrinsic horizon geometry can presumably be taken as another indication of ‘holographic’ properties of Lanczos-Lovelock models discussed in Sec. 2.7.

Two other aspects of Noether charge entropy worth mentioning in this context are the following. First, in Einstein gravity, we know that the York-Gibbons-Hawking boundary term, evaluated on the stretched horizon  $\mathcal{H}_\epsilon$ , also gives the correct horizon entropy in the limit  $\epsilon \rightarrow 0$ . A natural question, therefore, is whether the Lanczos-Lovelock surface term given in Eq. (56) also yields the same entropy. This definition of entropy was first given by Zanelli and Teitelboim [23], and an argument by Iyer and Wald in [41] shows that it matches with the Noether charge entropy. However, an explicit correspondence between these very different forms of the two expressions for entropy is not known in general. Perhaps the structural properties of the surface term described in detail in Sec. 3.2 and Sec. 3.3 might be useful to establish a direct correspondence between the two entropy expressions (at least for maximally symmetric horizons). (It is also possible to obtain the entropy of the horizon by a procedure (called ‘replica procedure’) related to the conical singularity and deficit angle in the Euclidean sector; this is discussed, e.g., in Ref. [45].)

Second, there has been extensive work, based on complex path formalism [48], to interpret the Hawking radiation as a tunnelling phenomenon (see e.g., [49, 50] and also [51]) and the question arises as to the status of this approach in Lanczos-Lovelock models. In this context, S. Sarkar and D. Kothawala were able to prove the following result for spherically symmetric horizons [52]: For any theory of gravity in which field equations near the horizon can be written in the form  $TdS - dE = PdV$ , the tunnelling formalism of [51] yields an entropy which matches the Wald entropy *once a specific choice is made for the Hamiltonian used in the tunnelling formalism*. As we shall show in section Sec. 4.1.2, the above criterion is satisfied near any static horizon (not necessarily spherically symmetric) in Lanczos-Lovelock models, with  $S$  being the Wald entropy (see Sec. 4.1.2 for definition of  $E$  etc.), and hence the result of [52] implies that the tunnelling mechanism can reproduce the correct expression for entropy in Lanczos-Lovelock models as a simple consequence of the first law of thermodynamics.

We conclude this section with a brief mention of an approach to the entropy of horizons, which does *not* seem to generalise from Einstein's theory to Lanczos-Lovelock models easily. This approach is based on entanglement of quantum fields. The modes of the quantum fields outside the horizon are entangled with modes inside and when the latter are traced out, the physical processes outside the horizon is governed by a density matrix  $\rho \propto \exp(-\beta H)$ . A natural entropy associated with this density matrix is  $S = -\text{Tr}(\rho \log \rho)$ . This quantity has a divergent behaviour of the form  $\lim_{\epsilon \rightarrow 0} (A/4\epsilon^2)$  where  $A$  is the horizon area and  $\epsilon$  has dimensions of length. If one cuts-off the limit at Planck length, one can match the expression with the standard result in Einstein gravity with entropy coming out as proportional to area. (This cut-off can be introduced in a covariant manner, using, for e.g., the techniques discussed in [46].) But since the entropy in Lanczos-Lovelock models is not, in general, proportional to area it is not clear whether the idea of entanglement entropy generalizes beyond Einstein's theory. (If it does not, it raises the possibility that it may not be the correct approach even in Einstein's theory.) The fact that it is divergent makes it difficult to draw firm conclusions. It is, for example, possible that the regularization procedure to introduce the cut-off somehow leads to the correct expression [47] but these ideas have not been worked out completely. So it is fair to say that whether entanglement entropy approach reproduces the correct Lanczos-Lovelock entropy is at present unclear.

### 3.2 Entropy of self-gravitating systems and Wald entropy

It is well known that statistical mechanics of systems with long range interactions exhibit various peculiar features (see, for a review, Ref. [53]) which could lead to subtleties in the thermodynamic description of such systems. It is therefore natural to ask whether effects of self-gravity can alter the conventional (extensive) nature of entropy of a matter configuration that is undergoing gravitational collapse, when the matter configuration is on the verge of forming a black hole.

This issue was first raised by Israel et. al. [54] in an attempt to give an operational definition of black hole entropy in Einstein gravity. Their results reproduced the well known area scaling of entropy in the limit of a matter shell collapsing to form a black hole (probably not a surprising result for a single shell), but more importantly, gave the correct numerical coefficient of  $1/4$  in the near horizon limit. This then motivated Oppenheim [55] to study a configuration of closely packed shells and ask how exactly the conventional volume scaling of entropy transmutes to an area scaling as the horizon is approached. The results obtained in

these works (also see [56]) suggested that entropy scaling approaches the Bekenstein-Hawking area law for self-gravitating systems provided the shells are assumed to be instantaneously in equilibrium with the local acceleration temperature in a manner made precise in the discussion below.

This result seems too good to be true at a fundamental level, since it seemingly provides an explanation for origin of black hole entropy without any appeal to quantum gravity! Perhaps this is just what happens, but one must be cautious since there are several ways of obtaining the correct entropy scaling using very different arguments. The contribution of effects of self-gravity to horizon entropy can be more firmly established if we could test the above analysis for theories other than Einstein gravity, since that would help isolate dynamics from kinematics in the final result. In this section, we shall describe the re-analysis of this approach for Lanczos-Lovelock models, and show that several non-trivial differences arise. The conclusion is that horizon entropy and entropy due to effects of self-gravity are different from each other in general, although the difference itself turns out to have an interesting form. This is one example in which the study of Lanczos-Lovelock models allows us to clearly isolate certain numerical coincidences which occur in Einstein's theory and do not carry over to Lanczos-Lovelock models.

To calculate the standard thermodynamic entropy of a self gravitating system, we will basically follow the set up suggested by Oppenheim [55], but in a slightly different manner to bring out the key features more prominently. We consider a system of  $n$  densely packed spherically symmetric shells in  $D$  spacetime dimensions, assumed to be in thermal equilibrium and supporting itself against its own gravity. We shall be interested in the entropy of this system when the outermost shell is close to the event horizon of the system. We first describe a general set-up to study the system described above for any spherically symmetric spacetime. This is important for two reasons. First, as we shall see, it will highlight the key *mathematical feature* responsible for entropy density  $S/A$  of the system being  $1/4$  for Einstein's theory in arbitrary dimensions. In [55], this result was obtained for  $D = 4$  by explicitly using the Schwarzschild form of metric functions and the expression for Hawking temperature. However, we shall not require any explicit expression for metric functions or Hawking temperature to obtain the area scaling. Second, the set-up we describe can be used, without any modification, to obtain entropy density for higher curvature gravity actions.

Let us denote the variables describing the  $i$ th shell with a subscript  $i$ . Since the system is spherically symmetric, we can write the metric outside the  $i$ th shell as

$$ds^2 = -c_i f_i(r) dt^2 + b_i(r)^{-1} dr^2 + r^2 d\Omega^2 \quad (101)$$

where  $d\Omega^2$  is the metric of a unit  $(D - 2)$  sphere. At the location of the shells (given by  $r_i = \text{const}$ ), the above metric must satisfy the Israel junction condition which states that the induced metric on the hypersurface should be continuous. This leads to the following conditions on the constants  $c_i$ :

$$c_i f_i(r_{i+1}) = c_{i+1} f_{i+1}(r_{i+1}) \quad (102)$$

and we shall choose  $c_n = 1$  (which fixes the interpretation of  $t$  as proper time for asymptotic observers). Then Eq. (102) can be solved to give

$$c_k = \frac{f_{k+1}(r_{k+1})}{f_k(r_{k+1})} \dots \frac{f_{n-1}(r_{n-1})}{f_{n-2}(r_{n-1})} \frac{f_n(r_n)}{f_{n-1}(r_n)} \quad (103)$$

Note that when  $f_n(r_n) = 0$  — i.e., when the outermost shell is exactly on the horizon — we have  $c_n = 0$  for all  $(i \neq n)$  which implies that

$$g_{00} = 0 \quad \forall i \quad (104)$$

Here the condition that the  $g_{00}$ 's vanish even for the inner-shells indicates that our assumption regarding the staticity of the inner shells is not valid when  $f_n(r_n)$  is exactly zero. (This is because we know that a particle cannot be kept at a fixed position inside a black hole.) However, for the purpose of our discussion, we only need to consider the limit in which the outer shell is very near to the horizon, that is,  $f_n(r_n) \rightarrow 0$ , then the result that the  $c_i$ 's (other than  $c_n$ ) and  $g_{00}$  vanish is just the leading order result in this approximation. Henceforth, we shall assume that we are working in this limit and will not state it explicitly unless otherwise needed.

We shall next use the second junction condition which relates jumps in geometric quantities across a shell, to the matter stress tensor of the shell,  $t_{(i)\mu\nu}$ . Let the normal to a  $r = \text{const.}$  surface (dropping the subscript  $i$  for convenience) be  $n_a = b^{-1/2} \partial_a r$ , so that the induced metric on  $r = \text{constant}$  surface becomes  $h_{\mu\nu} = (g_{ab} - n_a n_b) \delta_\mu^a \delta_\nu^b$  in coordinates  $(t, \theta_A)$ . Here, we will use Greek letters to define components in the  $r = \text{constant}$  surface with coordinates  $(t, \theta^A)$ . Within the hypersurface, we can further define  $u_\mu = \sqrt{cf} \partial_\mu t$ , and the induced metric on the level surfaces of  $t$ ,  $q_{\mu\nu} = h_{\mu\nu} + u_\mu u_\nu$ . It then follows from spherical symmetry that  $t_{\mu\nu}$  for each shell has the general form

$$t_{(i)\mu\nu} = \rho_i u_{(i)\mu} u_{(i)\nu} + P_i q_{(i)\mu\nu} \quad (105)$$

with  $E_i = 4\pi r_i^2 \rho_i$  is the energy of the shell. The physical interpretation of  $E_i$  and  $P_i$  is that of energy and pressure as measured by a local observer at rest on the shell. Further, the condition for thermal equilibrium in curved spacetime implies [57]

$$\begin{aligned} T_i \sqrt{-g_{00}^i(r_i)} &= T_\infty \\ \mu_i \sqrt{-g_{00}^i(r_i)} &= \mu_n \sqrt{-g_{00}^n(r_n)} \end{aligned} \quad (106)$$

where  $T_\infty$  is the temperature of the system as measured by a static observer at infinity, and  $\mu_i$ 's denote the chemical potential. We can thus express all the  $T_i$ 's and  $\mu_i$ 's in terms of just two unknown parameters  $T_\infty$  and  $\mu_n$  respectively. In thermodynamic equilibrium, each shell satisfies the Gibbs-Duhem relation

$$E_i = T_i S_i - P_i A_i + \mu_i N_i \quad (107)$$

where  $N_i$  is the number of particles composing the  $i$ th shell and  $A_i$  is the area of the  $i$ th shell. Using Eq. (106) in the above expression, we can write the total entropy  $S_{\text{matter}}$  of the



system in the form

$$\begin{aligned}
S_{\text{matter}} &= \sum_i S_i \\
&= \sum_i \frac{E_i + P_i A_i}{T_\infty} \sqrt{-g_{00}^i(r_i)} - \frac{\mu_n N}{T_\infty} \sqrt{-g_{00}^n(r_n)} \\
&= \sum_i \frac{E_i + P_i A_i}{T_\infty} \sqrt{c_i f_i(r_i)} - \frac{\mu_n N}{T_\infty} \sqrt{f_n(r_n)}
\end{aligned} \tag{108}$$

Now assuming that  $\mu_n N$  is a finite quantity, the last term in the above expression vanishes in the near horizon limit, since  $g_{00}^n(r_n)$  vanishes (see Eq. (104)). Hence the only non-trivial contribution to  $S_{\text{matter}}$  can come from the first term, which we wish to evaluate in the limit when outermost shell is close to the horizon of the system,  $r_n \rightarrow r_H$  (where  $f_n(r_H) = 0$ ); we shall refer to this as the near-horizon limit. We can read-off this contribution as follows.

First, we note that  $(E_i + P_i A_i)/T_\infty$  has no dependence on  $c_i$ , since  $t_{(i)\mu\nu}$  does not depend on  $c_i$  (see below). Further, in the near horizon limit,  $c_i = 0 \ \forall i \neq n, c_n = 1$ , from which it is easy to see that the only non-zero contribution to  $S_{\text{matter}}$  will come from the  $i = n$  term, and in particular from those terms in  $(E_n + P_n A_n)$  which diverge as  $1/\sqrt{f_n(r_n)}$ . Therefore, our strategy to calculate the entropy of the system will be to analyze the form of  $E_n$  and  $P_n$  and look for such divergent factors. As we shall now demonstrate, this strategy provides an extremely quick way to obtain the result in Einstein theory (compared to the explicit computations in [55]), and also facilitates easy generalization to higher derivative theories.

In Einstein theory, the surface stress tensor is given by

$$8\pi t_{(i)\mu\nu}^{(E)} = \left\langle K h_{\mu\nu} - K_{\mu\nu} \right\rangle_i \tag{109}$$

where  $\langle \dots \rangle_i = \langle \dots \rangle_{i+1} - \langle \dots \rangle_i$  evaluated at  $r = r_i$ . The extrinsic curvature of  $r = \text{const}$  surface, for the metrics described by Eq. (101), is given by (after dropping the subscript  $i$  for convenience):

$$\begin{aligned}
K_{\mu\nu} &= -\left(\sqrt{b}f'/2f\right) u_\mu u_\nu + \left(\sqrt{b}/r\right) q_{\mu\nu} \\
K &= h^{\mu\nu} K_{\mu\nu} \\
&= \left(\sqrt{b}f'/2f\right) + (D-2) \left(\sqrt{b}/r\right)
\end{aligned} \tag{110}$$

Note that, as stated above, there is no dependence on  $c_i$ . Based on our strategy outlined in the previous section, we must now find terms which, for  $i = n$ , diverge as  $1/\sqrt{f_n(r_n)}$  as  $r_n \rightarrow r_H$ . We now specialize to the case  $b_i(r) = f_i(r) \ \forall i$ . (This condition is automatically satisfied, in the near-horizon limit, for all the cases in which the horizon has no curvature singularity; so this choice entails no loss of generality.) Using Eqs. (109), (110) and (108), we

can immediately write, for the divergent parts,

$$\begin{aligned} 8\pi[\rho_n]_{\text{div}} &= 8\pi \left[ t_{(n)\hat{0}\hat{0}}^{(E)} \right]_{\text{div}} = 8\pi [t_{(n)\mu\nu} u^\mu u^\nu]_{\text{div}} = 0 \\ 8\pi[P_n]_{\text{div}} \delta_B^A &= 8\pi \left[ t_{(n)B}^{(E)A} \right]_{\text{div}} = -K_{\hat{0}\hat{0}}^{(n)} \delta_B^A \end{aligned} \quad (111)$$

in the limit  $r_n \rightarrow r_H$ ,  $f_n(r_H) \rightarrow 0$  (here  $K_{\hat{0}\hat{0}}^{(n)} = K_{\mu\nu}^{(n)} u^\mu u^\nu$ ). We have recalled the definition of  $\rho_i$ 's and  $P_i$ 's from Eq. (105). Hence, we see that the main divergence comes from

$$\begin{aligned} [P_n]_{\text{div}} &= -\left(\frac{1}{8\pi}\right) K_{\hat{0}\hat{0}}^{(n)} \\ &= \left(\frac{1}{16\pi}\right) \lim_{r_n \rightarrow r_H^+} \frac{f'_n(r_n)}{\sqrt{f_n(r_n)}} \\ &= \frac{T_\infty}{4} \lim_{\epsilon \rightarrow 0^+} \frac{1}{\sqrt{f_n(r_H + \epsilon)}} \end{aligned} \quad (112)$$

where we have used the definition of Hawking temperature for a system which has collapsed to a black hole. Using the above result in Eq. (108), we finally get

$$\begin{aligned} S_{\text{matter}} &\xrightarrow{r_n \rightarrow r_H} S_n \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{(1/4)T_\infty / \sqrt{f_n(r_H + \epsilon)}}{T_\infty} A_H \times \sqrt{f_n(r_H + \epsilon)} \\ &= \frac{A_H}{4} \end{aligned} \quad (113)$$

Hence, we see that the entropy of the configuration, as the outermost shell approaches the event horizon of the system, is dominated by the Bekenstein-Hawking entropy. The crucial factor of  $1/4$  comes from the last line of Eq. (112), which can be recast as

$$[P_n]_{\text{div}}/T_n = 1/4 \quad (114)$$

Interestingly, this same relation arises in the context of a derivation of Navier-Stokes equation and its relation to emergent gravity paradigm [58].

In what follows, we shall repeat the above analysis for Lanczos-Lovelock gravity, and show that the correspondence between shell entropy as calculated above, and the black hole entropy as given by Wald formula *does not always hold*. We will also point out that our result, nevertheless, has a curious interpretation in terms of the “equation of state” of the horizon degrees of freedom, and also has a very direct mathematical connection with the membrane paradigm for black hole horizons.

As discussed in Sec. 3, the surface stress tensor in the  $m^{\text{th}}$  order Lanczos-Lovelock theory is given by

$$\begin{aligned} 8\pi \left( t_{(m)\mu}^\nu \right)_i &= \frac{m!}{2^{m+1}} \alpha_m \left\langle \sum_{s=0}^{m-1} C_{s(m)} \left( \pi_{s(m)} \right)_\mu^\nu \right\rangle_i \\ \left( \pi_{s(m)} \right)_\mu^\nu &= \delta_{[\mu\mu_1 \dots \mu_{2m-1}]}^{\nu\nu_1 \dots \nu_{2m-1}} \hat{R}_{\nu_1\nu_2}^{\mu_1\mu_2} \dots \hat{R}_{\nu_{2s-1}\nu_{2s}}^{\mu_{2s-1}\mu_{2s}} K_{\nu_{2s+1}}^{\mu_{2s+1}} \dots K_{\nu_{2m-1}}^{\mu_{2m-1}} \end{aligned} \quad (115)$$

where the coefficients  $C_{s(m)}$  are given by

$$C_{s(m)} = \sum_{q=s}^{m-1} \frac{4^{m-q} q C_s (-2)^{q-s}}{q! (2m-2q-1)!!} \quad (116)$$

Here  $\hat{R}_{\mu\rho\sigma\nu}$  etc. indicates that these quantities are to be evaluated for the induced metric  $h_{\mu\nu}$ . To calculate the matter entropy of the system of  $n$  shells, we proceed in the manner similar as in the case of Einstein's gravity described in the previous section, that is, by finding  $\left[t_{(n)\mu}^\nu\right]_{\text{div}}$ . We note that, in the near horizon limit (and assuming that the intrinsic Riemann tensor of horizon surface is finite), the only divergence in the stress tensor will come from the  $K_0^{\hat{0}}$  component of the extrinsic curvature of the  $r = r_n$ th shell, which diverges as  $(1/\sqrt{f_n})$  as can be seen from Eq. (110). The transverse components  $K_B^A$  vanish as  $(\sqrt{f_n}/r)$ . Further, note that the determinant tensor has the property

$$\delta_{0\beta_1\beta_2\cdots\beta_n}^{0\alpha_1\alpha_2\cdots\alpha_n} = \delta_{B_1B_2\cdots B_n}^{A_1A_2\cdots A_n} \times \left( \delta_{A_1}^{\alpha_1} \delta_{\beta_1}^{B_1} \cdots \delta_{A_n}^{\alpha_n} \delta_{\beta_n}^{B_n} \right) \quad (117)$$

That is, the presence of 0 in each row of the determinant tensor forces all the other indices to take the values  $2, 3, \dots (D-1)$ . Thus, at most, we can have only one  $K_0^{\hat{0}}$  present in any term of Eq. (61) which shows that the maximum possible divergence is  $O(1/\sqrt{f_n})$  even though the surface stress tensor is a  $(2m-1)$ th degree polynomial in  $K$ . From the structure of Eq. (61), it is easy to see that only the  $s = (m-1)$  term gives this divergent contribution while all the other terms being of  $O(\sqrt{f})$  or higher vanish while from Eq. (117), it is again obvious that only the transverse component of the surface stress tensor will contribute. Thus we get

$$\begin{aligned} 8\pi[\rho_n]_{\text{div}} &= 8\pi \left[ t_{(n)00} \right]_{\text{div}} = 8\pi \left[ t_{(n)\mu\nu} u^\mu u^\nu \right]_{\text{div}} = 0 \\ 8\pi[P_n]_{\text{div}} \delta_B^A &= 8\pi \left[ t_{(n)B}^A \right]_{\text{div}} = \frac{m\alpha_m}{2^{m-1}} \delta_{BB_1\cdots B_{2m-2}}^{AA_1\cdots A_{2m-2}} \hat{R}_{A_1A_2}^{B_1B_2} \cdots \hat{R}_{A_{2m-3}A_{2m-2}}^{B_{2m-3}B_{2m-2}} K_0^{\hat{0}} \end{aligned} \quad (118)$$

where in the second line we have used  $C_{m-1} = 4/(m-1)!$  from its definition in Eq. (58), and replaced  $\hat{R}_{CD}^{AB}$  with the intrinsic curvature  $\hat{R}_{CD}^{AB}$  defined completely in terms of the induced metric  $q_{AB}$  of the constant  $t$ , constant  $r$  surface. Such a replacement is valid since the corresponding extrinsic curvature vanishes. We can now obtain the pressure easily using the fact that  $q_{AB}$  is maximally symmetric (see Eq. (105)), and hence

$$\begin{aligned} [P_n]_{\text{div}} &= \frac{1}{(D-2)} q^{AB} \left[ t_{(n)AB} \right]_{\text{div}} = \frac{2m\alpha_m}{(D-2)} \frac{1}{16\pi} \frac{1}{2^{m-1}} \delta_{AB_1\cdots B_{2m-2}}^{AA_1\cdots A_{2m-2}} \hat{R}_{A_1A_2}^{B_1B_2} \cdots \hat{R}_{A_{2m-3}A_{2m-2}}^{B_{2m-3}B_{2m-2}} K_0^{\hat{0}} \\ &= \left( \frac{D-2m}{D-2} \right) 2m\alpha_m L_{m-1}^{(D-2)} K_0^{\hat{0}} \end{aligned} \quad (119)$$

In arriving at the second equality, we have used the following two relations concerning Lanczos-Lovelock actions of order  $m$  in  $D$  dimensions derived earlier in Eqs. (30), (37)

$$\begin{aligned} E_{j(m)}^i &= -\frac{1}{2} \frac{1}{16\pi} \frac{1}{2^m} \delta_{jc_1d_1\cdots c_md_m}^{ia_1b_1\cdots a_mb_m} R_{a_1b_1}^{c_1d_1} \cdots R_{a_mb_m}^{c_md_m} \\ E_{i(m)}^i &= -\frac{D-2m}{2} L_m^D \end{aligned} \quad (120)$$

where  $E_{j(m)}^i$  is the equation of motion tensor for Lanczos-Lovelock action. Hence, we once again find that the only non-zero contribution to matter entropy comes from the pressure term as in the case of Einstein gravity (see Eq. (111)). No further computation is needed since we can simply obtain the entropy by comparing expression in Eq. (119) with the first of Eqs. (112) in Einstein theory. This immediately allows us to read-off the matter entropy as

$$S_{\text{matter}}^{(m)} = \left( \frac{D-2m}{D-2} \right) 4\pi m \alpha_m L_{m-1}^{(D-2)} A_H \quad (121)$$

where  $A_H$  is the  $D-2$  dimensional hypersurface area of the horizon. On the other hand, the Wald entropy for general Lanczos-Lovelock theory obtained earlier in Eq. (100), when evaluated for the spherically symmetric case is given by:

$$S_{\text{Wald}}^{(m)} = 4\pi m \alpha_m L_{m-1}^{(D-2)} A_H \quad (122)$$

which leads to the following relation

$$S_{\text{matter}}^{(m)} = \left( \frac{D-2m}{D-2} \right) S_{\text{Wald}}^{(m)} \quad (123)$$

So the matter entropy in Lanczos-Lovelock theory is proportional to the corresponding Wald entropy of the black hole. We also see that  $S_{\text{matter}} \leq S_{\text{Wald}}$  with the equality holding only in Einstein's gravity  $m=1$ . This inequality strongly suggests that, in general, the microscopic degrees of freedom responsible for the entropy of the black hole are certainly not the extrinsic (matter) degrees of freedom forming the black hole. Further, since a general Lanczos-Lovelock theory will be described by Lagrangian of the form  $L = \sum_m \alpha_m L_m$ , the total matter entropy in such theories would be

$$S_{\text{matter}} = \sum_{m=0}^{[(D-1)/2]} \left( \frac{D-2m}{D-2} \right) S_{\text{Wald}}^{(m)} \quad (124)$$

which has no simple relation, or proportionality to the Wald entropy

$$S_{\text{Wald}} = \sum_{m=0}^{[(D-1)/2]} S_{\text{Wald}}^{(m)} \quad (125)$$

We highlight another point which is relevant to the membrane paradigm for Lanczos-Lovelock models, to be discussed in Sec. 3.3. As is evident from above analysis, it is only the pressure term which leads to a non-zero contribution to the matter entropy and hence, in the near horizon limit, we have the result:

$$\begin{aligned} \frac{S_{\text{matter}}}{A_H} &= \frac{[P_n]_{\text{div}}}{T_n} \\ &\equiv \frac{P}{T_n} \end{aligned} \quad (126)$$

where for the sake of brevity, we have replaced  $[P_n]_{\text{div}}$  by  $P$ . Thus the scaling of entropy with the horizon radius  $r_H$ , as well as the proportionality constant, is completely determined

by the relation between  $P$  and  $T$ , which is to say that the equation of state  $P = P(T)$  determines the matter entropy. In the case of Lanczos-Lovelock gravity, we can read off the equation of state  $P = P(T) = [t_{(n)\theta}^\theta]_{\text{div}}$  from Eq. (119) by using the relation  $K_{\hat{0}\hat{0}} = -K_{\hat{0}}^{\hat{0}} = -f'_H/(2\sqrt{f}) = -2\pi T_n$  to get

$$\frac{P}{T_n} = \left( \frac{D-2m}{D-2} \right) 4\pi m \alpha_m L_{m-1}^{(D-2)} \quad (127)$$

Using Eq. (123), the above relation can also be written in terms of Wald entropy as:

$$\frac{PA_H}{T_n} = \left( \frac{D-2m}{D-2} \right) S_{\text{Wald}}^{(m)} \quad (128)$$

In fact, the pressure  $P$  obtained above is exactly what one would call as surface tension of the stretched horizon in the membrane paradigm. This is mathematically obvious from the way we arrived at the near-horizon for  $P$ , but it must be emphasized that the equivalence with membrane pressure is strictly valid only in the near-horizon limit.

The main implication of the above result is that matter entropy of a system is not necessarily the only contribution to entropy of a black hole formed from collapse of that system. This unambiguously establishes the fact that there might be more to horizon entropy than just the effects of self-gravitation, at least to the extent that the approximations made in the above analysis are valid. (We have not studied the system in full generality, since the dynamics of the collapse has not been taken into account. We only considered the shells at various stages of collapse, imposing consistency conditions in the form of junction conditions at each instant of the collapse. While this may be sufficient, it will be interesting to study what happens in a fully dynamical situation.) However, whatever be the nature of additional corrections introduced by a more general analysis, it is unlikely to establish the complete equivalence of the resultant entropy with the Wald entropy.

It is also interesting to view the above result in the light of several entropy bounds on matter entropy which exist in the literature. The model calculation above gives precise expressions for matter entropy and its relation to Wald entropy, rather than just an inequality. Finally, the analysis above also highlights a possible connection between horizon entropy and the “equation-of-state”  $P(T)$  satisfied by the degrees of freedom residing on the horizon surface. This would become more evident when we present a study of the *membrane paradigm* for the Lanczos-Lovelock models in the next section, deriving all the transport coefficients to linear order in the perturbation theory.

### 3.3 Membrane paradigm in Lanczos-Lovelock models

We now use the mathematical tools of the previous section to derive several properties of the so called stretched horizon in a theory described by the Lanczos-Lovelock action. That is, we set up the so called *membrane paradigm* for the Lanczos-Lovelock models. After a brief introduction to the membrane paradigm in Einstein gravity, we shall discuss the analysis for the Lanczos-Lovelock models, keeping in focus the new features introduced by the non-triviality of the Lanczos-Lovelock action.

Specifically, we shall present a derivation of the various *transport coefficients* associated with the membrane to first order in perturbation theory, and demonstrate how the structure

of these coefficients brings into focus several fundamental features of the membrane paradigm itself, which are otherwise not evident when attention is restricted solely to Einstein gravity. This reinforces our original point of view about the utility of Lanczos-Lovelock models in gaining insight into aspects of Einstein gravity itself.

The membrane paradigm of black holes [59] takes the point of view that, as far as the static observers outside the horizon are concerned, the black hole horizon can be replaced by a *stretched* horizon, a membrane, endowed with specific physical properties which encode the presence of the inaccessible black hole region. Although originally developed to facilitate inclusion of the physics of black hole background in astrophysical applications, work over the last decade or so has indicated that physical properties of the membrane (such as viscosity etc.) might have a deeper relevance, from the point of view of holographic dualities which map gravitational systems to non-gravitational systems in one lower dimension. This has led to a renewed interest in the study of membrane paradigm.

The membrane paradigm in Einstein gravity has also been analysed from a view point of the emergent gravity paradigm. The Damour-Navier Stokes equation governing the dynamics of the black hole membrane could be obtained [58] starting from an action which could be given a thermodynamic interpretation as an entropy production rate when expressed in terms of thermodynamic variables such as temperature, entropy, pressure, etc of the horizon. This comes very close to quantifying a precise connection of properties of the horizon membrane with the thermodynamics of horizons, but as we shall show in this section, a more direct connection between the transport coefficients of the membrane and its thermodynamic properties can be established by studying the Lanczos-Lovelock models.

In Einstein gravity (as well as with the Gauss-Bonnet term added to it), the *geometric* structure of the fluid transport coefficients of the membrane is not evident from their *algebraic* expressions. For example, in Einstein's gravity, it is known that the shear viscosity  $\eta$  is equal to  $1/(16\pi)$ . But the origin of this factor is not well understood in terms of the *geometric* properties of the horizon such as the intrinsic curvature of the horizon etc. As we shall show, setting up the membrane paradigm [60] for the Lanczos-Lovelock models once again helps us in not only understanding such geometric aspects of the transport coefficients in a more general manner, but even explains the origin of factors such as  $1/(16\pi)$  for shear viscosity in Einstein theory!

Let us briefly describe the geometric set-up and the procedure to construct the membrane paradigm. More details can be found in [59], and we only refresh the essential expressions below to set up the notation etc. We shall work in a  $D$  dimensional spacetime containing an event horizon  $H$  generated by the null geodesics  $\mathbf{l}$ . The surface gravity  $\kappa$  is then determined through  $\nabla_{\mathbf{l}}\mathbf{l} = \kappa\mathbf{l}$ . We shall deal with a stationary background spacetime, in which case  $\mathbf{l}$  corresponds to the timelike Killing vector  $\boldsymbol{\xi}$  restricted to  $H$  where  $\xi^2 = 0$ . For asymptotically flat spacetimes, the norm of  $\mathbf{l}$  can be fixed at infinity and there is no ambiguity in the definition of  $\kappa$ . The stretched horizon  $\mathcal{H}_s$  is then defined as a timelike surface infinitesimally close to  $H$ , and separated from  $H$  along spacelike geodesics  $\mathbf{n}$ . This surface is spanned by a unit timelike vector field  $\mathbf{u}$  and  $(D-2)$  vectors  $\mathbf{e}_{(A)}$ 's. The correspondence between points on  $H$  and  $\mathcal{H}_s$  is made via ingoing null geodesics  $\mathbf{k}$  normalized to have unit Killing energy,  $\mathbf{k} \cdot \mathbf{l} = -1$ . The horizon  $H$  is then located at  $\lambda = 0$ , where  $\lambda$  is the affine parameter along  $\mathbf{k}$ , related to the norm of  $\boldsymbol{\xi}$  by  $\lambda = -(1/2\kappa)\xi^2$ . We shall define  $N = \sqrt{-\xi^2}$  to simplify notation and also to facilitate comparison with known results in the literature. In the limit  $\lambda \rightarrow 0$ ,  $\mathcal{H}_s \rightarrow H$  and  $\mathbf{u} \rightarrow N^{-1}\mathbf{l}$  and  $\mathbf{n} \rightarrow N^{-1}\mathbf{l}$ .

The induced metric  $h_{\mu\nu}$  on  $\mathcal{H}_s$  is given by  $h_{\mu\nu} = h_{ab}e_{(\mu)}^a e_{(\nu)}^b$ , where  $h_{ab} = g_{ab} - n_a n_b$  and  $e_{(\mu)}$ 's are basis vectors spanning  $\mathcal{H}_s$  ( $e_{(\mu)} \cdot \mathbf{n} = 0$ ). Similarly, the induced metric  $\gamma_{AB}$  on the  $(D-2)$ -dimensional space-like cross-section of  $\mathcal{H}_s$  orthogonal to  $\mathbf{u}$  is given by  $\gamma_{AB} = \gamma_{ab}e_{(A)}^a e_{(B)}^b$  where  $\gamma_{ab} = h_{ab} + u_a u_b$  and  $e_{(A)}$ 's are basis vectors spanning  $\gamma$  ( $e_{(A)} \cdot \mathbf{u} = 0 = e_{(A)} \cdot \mathbf{n}$ ). The extrinsic curvature of  $\mathcal{H}_s$  is defined as  $K_{\mu\nu} = e_{(\mu)}^a e_{(\nu)}^b \nabla_a n_b$  and it is easy to verify that in the limit  $N \rightarrow 0$ , we have:

$$\begin{aligned} K_{\hat{0}\hat{0}} &= K_{\mu\nu} u^\mu u^\nu = -N^{-1} \kappa \\ K_{\hat{0}A} &= K_{\mu\nu} u^\mu e_{(A)}^\nu = 0 \\ K_{AB} &= K_{\mu\nu} e_{(A)}^\mu e_{(B)}^\nu \\ &= N^{-1} k_{AB}, \end{aligned} \tag{129}$$

where  $k_{AB}$  is the extrinsic curvature of the  $(D-2)$ -dimensional space-like cross-section of the true horizon  $H$ . This can be decomposed with respect to its trace-free part as

$$k_{AB} = \sigma_{AB} + \frac{1}{(D-2)} \theta \gamma_{AB}, \tag{130}$$

where  $\gamma^{AB} \sigma_{AB} = 0$ , and  $\theta$  and  $\sigma_{AB}$  are the expansion scalar and shear respectively of the  $l$  congruence generating the horizon. For a Killing horizon, both of these vanish as  $N^2$  and therefore  $K_{\hat{0}\hat{0}}$  remains the only divergent contribution to  $K_{\mu\nu}$ . However, since we wish to investigate dissipative properties associated with the horizon, we shall perturb the background away from staticity and then obtain the membrane stress tensor to first order in perturbations to determine the transport coefficients. The extrinsic curvature can be written, upto linear order in perturbations, as

$$K_\beta^\alpha = K_\beta^{(0)\alpha} + K_\beta^{(1)\alpha} \tag{131}$$

where  $K_\beta^{(0)\alpha}$  is the unperturbed extrinsic curvature. Since we assumed the background geometry to be static, we have

$$K_{\hat{0}\hat{0}}^{(0)} = -\kappa/N, \quad K_{\hat{0}A}^{(0)} = 0, \quad K_{AB}^{(0)} = (N/r) \delta_{AB} \tag{132}$$

To study the additional divergences brought in when the horizon is perturbed, it is convenient to choose perturbations which do not affect  $K_{\hat{0}\hat{0}}^{(0)}$ . (As we shall see, our final relations will only have divergence of  $O(N^{-1})$ , and perturbing  $\kappa$  will not lead to any qualitative difference.) Formally, since we assume the background  $\kappa$  to be non-zero, the perturbation in  $\kappa$  is not gauge invariant and our choice corresponds [59] to a gauge wherein  $\delta\kappa = 0$ .) Therefore, we have

$$K_{\hat{0}\hat{0}}^{(1)} = 0, \quad K_{\hat{0}A}^{(1)} = 0, \quad K_{AB}^{(1)} = (1/N) k_{AB} \tag{133}$$

where we have not bothered to put a “(1)” over  $k_{AB}$  since its unperturbed part is strictly zero. The expansion and shear parameters of the perturbed horizon will now introduce additional divergences in the stress tensor via  $K_{AB}$ .

We are now ready to use the above facts to analyze the membrane stress tensor in Lanczos-Lovelock theories of gravity. For further details about the membrane paradigm and aspects of the perturbation scheme, we refer the reader to earlier literature on the subject [59, 61], and also to the introductory sections in the recent paper [62] which focuses on Einstein-Gauss-Bonnet theories. Since the derivation for Lanczos-Lovelock involves considerable combinatorics, we present a summary of the final results first and then provide a sketch of the derivation [60].

### 3.3.1 Membrane paradigm: summary of the results

We shall be mainly interested in membranes which have isotropic tangential stresses (pressure), which is only possible when all directions, everywhere within the horizon surface, are equivalent. (That is, the horizon surface is maximally symmetric). Therefore, the cross-section of the horizon of the background geometry has the curvature

$${}^{(D-2)}R_{ABCD} = (\mathcal{K}/r^2) (\gamma_{AC}\gamma_{BD} - \gamma_{AD}\gamma_{BC}) \quad (134)$$

where  $\mathcal{K} = \pm 1$  and  $r$  is a constant of dimension *length* which is related to the intrinsic Ricci scalar of the horizon cross-section as  $\mathcal{K}(D-3)(D-2)r^{-2} = {}^{(D-2)}R$ . For simplicity of notation, we will assume  $\mathcal{K} = +1$  in the intermediate steps, and only restore it in the final expressions (which is easy to do). Doing so will come in handy while discussing the case of planar horizons, for which we will simply put  $\mathcal{K} = 0$ . With this assumption, it can be shown that the horizon membrane has the stress tensor

$$t_{\alpha\beta} = \rho_s u_\alpha u_\beta + e_\alpha^{(A)} e_\beta^{(B)} (p_s \gamma_{AB} - 2\eta_s \sigma_{sAB} - \zeta_s \theta_s \gamma_{AB}) \quad (135)$$

with  $p_s = \sum_m p_s^{(m)}$  etc., and the different variables are given by:

$$\begin{aligned} \text{Pressure : } p_s^{(m)} &= \left( \frac{D_{2m}}{D_2} \right) \left( \frac{\kappa}{2\pi} \right) 4\pi m \alpha_m L_{m-1}^{D-2} \\ \text{Energy density : } \rho_s^{(m)} &= - \left( \frac{\theta_s}{\kappa} \right) p_s^{(m)} \\ \text{Shear Viscosity : } \eta_s^{(m)} &= \left( \frac{m \alpha_m r_H^2}{\mathcal{K} D_2 D_3} \right) \left[ L_m^{D-2} - \left( \frac{2\kappa}{r_H} \right) (m-1) D_{2m} L_{m-1}^{D-2} \right] \\ \text{Bulk Viscosity : } \zeta_s^{(m)} &= - \left( \frac{2D_3}{D_2} \right) \eta_s^{(m)} \end{aligned} \quad (136)$$

where we have introduced the notation  $D_k \equiv (D-k)$  to avoid clutter, and  $\alpha_m$  is the coupling constant of the  $m^{\text{th}}$  order Lanczos-Lovelock Lagrangian

$$L_m^D = \frac{1}{16\pi} \frac{1}{2^m} \delta_{c_1 d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} \quad (137)$$

The transport coefficients above can be connected with the thermodynamic properties of the horizon in the Lanczos-Lovelock gravity, by noting that the Wald entropy and quasi-local



energy of the horizon in these theories are given by

$$\begin{aligned} S_{\text{Wald}}^{(m)} &= 4\pi m \alpha_m \int_H d\Sigma L_{m-1}^{D-2} \\ E_{\text{Wald}}^{(m)} &= \alpha_m \int^\lambda d\lambda \int_H d\Sigma L_m^{D-2} \end{aligned} \quad (138)$$

This immediately implies that

$$\frac{p_s}{T_\infty} \stackrel{N \rightarrow 0}{\equiv} \sum_{m=0}^{[D-1]/2} \left( \frac{D-2m}{D-2} \right) S_{\text{Wald}}^{(m)} \quad (139)$$

Similarly, it is obvious from the dependence on  $L_{m-1}^{D-2}$  and  $L_m^{D-2}$  factors appearing in the third and fourth equalities in Eqs. (136) for the shear and bulk viscosities, that both  $\eta$  and  $\zeta$  are also expressible solely in terms of  $S^{(m)}$  and  $E^{(m)}$ .

### 3.3.2 Membrane paradigm: details of the analysis

As discussed in Sec. 2.6, for the  $m^{\text{th}}$  order Lanczos-Lovelock Lagrangian  $L_m$  in Eq. (20), the surface stress tensor is given by (see Eq. (61)):

$$\begin{aligned} 8\pi t_{(m)\mu}^\nu &= \frac{\alpha_m m!}{2^{m+1}} \sum_{s=0}^{m-1} C_s \pi_{(s)\mu}^\nu \\ \pi_{(s)\mu}^\nu &= \delta_{[\mu\mu_1 \dots \mu_{2m-1}]}^{[\nu\nu_1 \dots \nu_{2m-1}]} \hat{R}_{\nu_1\nu_2}^{\mu_1\mu_2} \dots \hat{R}_{\nu_{2s-1}\nu_{2s}}^{\mu_{2s-1}\mu_{2s}} K_{\nu_{2s+1}}^{\mu_{2s+1}} \dots K_{\nu_{2m-1}}^{\mu_{2m-1}} \end{aligned} \quad (140)$$

with

$$C_s = \sum_{q=s}^{m-1} \frac{(-2)^{q-s} 4^{m-q} \binom{q}{s}}{q! (2m-2q-1)!!} \quad (141)$$

[We have replaced  $c_m$  in Eq. (61) by  $\alpha_m$  for notational clarity.] One can check that for  $m=1$  and  $m=2$ , the above expression reduces to that of the surface stress tensor in Einstein [26] and Gauss-Bonnet [27] theories respectively.

As in the case of the membrane paradigm for Einstein gravity, we will interpret the  $t_{\alpha\beta}$  as due to a fictitious matter source residing on the stretched horizon  $\mathcal{H}_s$ . Using the assumptions and expressions given above we can show that the surface stress energy tensor can be written in a form analogous to that of a viscous fluid, namely,

$$t_{\alpha\beta} = \rho_s u_\alpha u_\beta + e_\alpha^{(A)} e_\beta^{(B)} (p_s \gamma_{AB} - 2\eta_s \sigma_{sAB} - \zeta_s \theta_s \gamma_{AB}) \quad (142)$$

and thus read off the corresponding energy density  $\rho_s$ , pressure  $p_s$ , shear viscosity co-efficient  $\eta$  and co-efficient bulk viscosity  $\zeta_s$  in terms of the geometrical quantities describing the horizon.

From the fact that  $\pi_{(s)\mu}^\nu$  in Eq. (61) is a polynomial of odd degree  $[2(m-s)-1]$  in  $K_{\mu\nu}$ , it is easy to see that, to first order in perturbations, the only divergences in  $t_{\alpha\beta}$  are of  $O(N^{-1})$ , and they arise from  $s=(m-1)$  and  $s=(m-2)$  terms in the series. This follows (upon simple counting) from Eqs. (132), (133), (61) and (117). All the other terms in the series are

of  $O(N)$  and hence vanish in the limit  $N \rightarrow 0$ . We give below an outline of the contribution of divergent terms, stating only the relevant expressions and skipping the intermediate steps which are algebraically straightforward. We will also define

$$\left(Q_m^{(D-1)}\right)^{\alpha\mu\beta\nu} = (1/m) \partial L_m^{(D-1)} / \partial \hat{R}_{\alpha\mu\beta\nu} \quad (143)$$

A. *Divergence due to the zeroth order term:*

Only the  $s = (m-1)$  term will contribute to the zeroth perturbative order. The corresponding contribution [60] turns out to be

$$\begin{aligned} \text{For } s = (m-1) : \quad \left[\pi_{(s)\hat{0}\hat{0}}\right]_{\text{div}} &= 0 \\ \left[\pi_{(s)B}^A\right]_{\text{div}} &= N^{-1} \times 2^m 16\pi \left(Q_m^{(D-1)}\right)^{\hat{0}A}_{\hat{0}B} \kappa \end{aligned} \quad (144)$$

B. *Divergence due to the first order term:*

At first order in the perturbations, we get divergences from  $s = (m-1)$  and  $s = (m-2)$ . The contributions from these can also be found in a straightforward way, and are given by

$$\begin{aligned} \text{For } s = (m-1) : \quad \left[\pi_{(s)\hat{0}\hat{0}}\right]_{\text{div}} &= N^{-1} \times 2^m 16\pi \left(Q_m^{(D-1)}\right)^{\hat{0}C}_{\hat{0}D} k_C^D \\ \left[\pi_{(s)B}^A\right]_{\text{div}} &= N^{-1} \times 2^m 16\pi \left(Q_m^{(D-1)}\right)^{AC}_{BD} k_C^D \\ \text{For } s = (m-2) : \quad \left[\pi_{(s)\hat{0}\hat{0}}\right]_{\text{div}} &= 0 \\ \left[\pi_{(s)B}^A\right]_{\text{div}} &= 6 N^{-1} \times \frac{\kappa}{r} \times \delta_{\hat{0}FBD}^{\hat{0}FAC} \underbrace{\begin{matrix} \bullet\bullet\cdots\bullet\bullet \\ \bullet\bullet\cdots\bullet\bullet \end{matrix}}_{2(m-2)\text{factors}} \hat{R}_{\bullet\bullet} \cdots \hat{R}_{\bullet\bullet}^{\bullet\bullet} k_C^D \end{aligned} \quad (145)$$

where the symmetry factor of 6 in the last equation comes because we need to choose, from 3 factors of  $K_{\alpha\beta}$ , one factor of  $\overset{(1)}{K}_{AB}$  (3 ways) and one factor of  $K_{uu}$  from the remaining 2  $K_{\alpha\beta}$ 's (2 ways), which can be done in a total of  $3 \times 2 = 6$  ways.

The above equations can be further simplified by using the following identities (the proofs for which are given in Appendix C)

$$\left(Q_m^{(D-1)}\right)^{0A}_{0B} = \left(\frac{D_{2m}}{2D_2}\right) L_{m-1}^{D-2} \delta_B^A \quad (146)$$

$$\left(Q_m^{(D-1)}\right)^{AC}_{BD} = \frac{1}{2} \left(\frac{L_m^{D-2}}{D-2R}\right) \delta_{BD}^{AC} \quad (147)$$

$$\delta_{B\hat{0}DF}^{A\hat{0}CE} \underbrace{\begin{matrix} \bullet\bullet\cdots\bullet\bullet \\ \bullet\bullet\cdots\bullet\bullet \end{matrix}}_{2(m-2)\text{factors}} \hat{R}_{\bullet\bullet} \cdots \hat{R}_{\bullet\bullet}^{\bullet\bullet} = \left(\frac{2^{m-2}D_{2m}}{D_4}\right) \left(\frac{16\pi L_{m-1}^{D-2}}{D-2R}\right) \delta_{BDF}^{ACE} \quad (148)$$

where we have again used the notation  $D_k = (D - k)$  introduced earlier. We can now collect the  $O(N^{-1})$  terms in the total stress tensor  $t_{\alpha\beta} = t_{\alpha\beta}^{(0)} + t_{\alpha\beta}^{(1)}$  upto first order in perturbation to get

$$t_0^{\hat{0}} = N^{-1} \times 2m\alpha_m \left( \frac{D_{2m}}{D_2} \right) L_{m-1}^{D-2} \theta \quad (149)$$

$$t_B^A = N^{-1} \times 2m\alpha_m \left\{ \left( \frac{D_{2m}}{D_2} \right) L_{m-1}^{D-2} \kappa \delta_B^A + \left( \frac{L_m^{D-2}}{D-2 R} \right) k_B^A - 2(m-1)D_{2m} \left( \frac{L_m^{D-2}}{D-2 R} \right) \frac{\kappa}{r_H} k_B^A \right\} \quad (150)$$

We have thus accomplished our main task of finding the divergent contributions to surface stress tensor  $t_{\mu\nu}$  of the membrane to linear order in horizon perturbations. The only divergence is of  $O(N^{-1})$ , which is precisely the same as in Einstein theory, and can therefore be regulated in the same way as is done in conventional membrane paradigm for Einstein gravity. That is, we simply multiply by  $N$  to obtain the finite part of  $t_{\mu\nu}$ , with the regularization having the interpretation of canceling the infinite redshift at the horizon. Writing  $k_B^A$  in terms of the traceless part  $\sigma_B^A$  and trace  $\theta$  as in Eq. (130), we can express the stress tensor in the required form

$$t_{\alpha\beta} = \rho_s^{(m)} u_\alpha u_\beta + e_\alpha^{(A)} e_\beta^{(B)} \left( p_s^{(m)} \gamma_{AB} - 2\eta_s^{(m)} \sigma_{sAB} - \zeta_s^{(m)} \theta_s \gamma_{AB} \right) \quad (151)$$

and read off the pressure, transport co-efficients, etc as

$$\begin{aligned} \text{Pressure : } p_s^{(m)} &= \left( \frac{D_{2m}}{D_2} \right) \left( \frac{\kappa}{2\pi} \right) 4\pi m\alpha_m L_{m-1}^{D-2} \\ \text{Energy density : } \rho_s^{(m)} &= - \left( \frac{\theta_s}{\kappa} \right) p_s^{(m)} \\ \text{Shear Viscosity : } \eta_s^{(m)} &= \left( \frac{m\alpha_m}{(D-2)R} \right) \left[ L_m^{D-2} - \left( \frac{2\kappa}{r_H} \right) (m-1)D_{2m} L_{m-1}^{D-2} \right] \\ \text{Bulk Viscosity : } \zeta_s^{(m)} &= - \left( \frac{2D_3}{D_2} \right) \eta_s^{(m)} \end{aligned} \quad (152)$$

where, again,  $D_k \equiv (D - k)$ . One can check that for Einstein's gravity  $m = 1$  and Gauss-Bonnet gravity  $m = 2$ , the above expressions reduce to the corresponding expressions in the literature [59, 62]. But of course, as is evident, the final expressions yield much more than just a generalization of conventional results from Einstein theory to Lanczos-Lovelock models. As we already indicated earlier the above expressions highlight a connection between the fluid transport coefficients and thermal properties of the horizon and calls for a more thorough investigation into possible physical implications of the results.

In particular, one quantity which has gained significance from the point of view of the AdS-CFT correspondence is the ratio of shear viscosity to entropy density,  $\eta/s$ . Using the

expression of Wald entropy given in Eq. (138) in above expressions, we arrive at

$$\begin{aligned}
\frac{\eta^{(m)}}{s^{(m)}} &= \left( \frac{1}{4\pi^{D-2}R} \right) \left[ \frac{L_m^{D-2}}{L_{m-1}^{D-2}} - \left( \frac{2\kappa}{r_H} \right) (m-1) D_{2m} \right] \\
&= \left( \frac{D_{2m}}{4\pi D_2 D_3} \right) \left[ D_{2m+1} - 2(\kappa r_H)(m-1)\mathcal{K}^{-1} \right] \\
&\equiv \frac{C_m}{4\pi}
\end{aligned} \tag{153}$$

where we have used the relation  $L_m^{D-2} = (\mathcal{K}/r^2) D_{2m} D_{2m+1} L_{m-1}^{D-2}$  which holds for a  $(D-2)$  dimensional maximally symmetric spacetime. We can therefore write

$$\frac{\eta}{s} = \frac{1}{4\pi} \left[ \frac{1 + \sum_{m=2}^{m_c} C_m \bar{s}_m}{1 + \sum_{m=2}^{m_c} \bar{s}_m} \right] \tag{154}$$

where  $m_c = [(D-1)/2]$  and  $\bar{s}_m := s_m/s_1 = 4s_m$ .

One can now investigate the ratio  $\eta/s$  on a case-by-case basis, considering specific solutions with horizons in Lanczos-Lovelock theory. Specifically, for *planar*<sup>1</sup> horizons,  $\mathcal{K} = 0$ , and we note that in this case,  $C_m \rightarrow (\dots)\mathcal{K}^{-1}$ . Therefore, since  $\bar{s}_m \rightarrow (\dots)\mathcal{K}^{m-1}$ , only the  $m = 2$  term in the sum in the numerator survives for  $\mathcal{K} = 0$ . One is therefore led to the conclusion that no Lovelock term other than Gauss-Bonnet contributes to the  $\eta/s$  ratio for the planar ( $\mathcal{K} = 0$ ) case; this contribution is given by

$$\begin{aligned}
\text{For } \mathcal{K} = 0 : \quad \frac{\eta}{s} &= \frac{1 + [C_2 \bar{s}_2]_{\mathcal{K}=0}}{4\pi} \\
&= \frac{1}{4\pi} \left[ 1 - 4\alpha_2(D-4) \left( \frac{\kappa}{r_H} \right) \right]
\end{aligned} \tag{155}$$

Although only the Gauss-Bonnet contributes for the planar case, it must be noted that  $\eta/s$  will, in general, depend on other Lovelock coupling constants as well, through the ratio  $(\kappa/r_H)$  which must be calculated using a specific  $\mathcal{K} = 0$  solution of the full Lovelock action.

A similar result has been noted in the literature [63, 64] in the context of AdS-CFT correspondence with the difference being in the boundary used — which is the AdS infinity whereas we have considered the boundary to be the stretched horizon. However, it has been argued in [65] that the  $\eta/s$  ratio for the two cases are related by a renormalization-group flow equation in Einstein's gravity. It would indeed be interesting to check whether such a flow exist in Lanczos-Lovelock theory, since our result for the membrane paradigm, in conjunction with the AdS-CFT results for Lanczos-Lovelock gravity, does indicate such a

---

<sup>1</sup>A couple of clarifying points while dealing with planar horizons are in order: (i) although  $(\kappa r_H)$  would in general depend on  $\mathcal{K}$ , we expect it to be finite for  $\mathcal{K} = 0$ , and (ii) the most general form of a planar horizon metric, consistent with the assumption of staticity, is:  $ds_H^2 = \Omega(r/L_i) (dx^2 + dy^2 + \dots)$ , where  $\Omega(r/L_i)$  is a dimensionless function constructed out of ratios of the coordinate  $r$  and any other length scale(s)  $L_i$  appearing in the metric. For known solutions which are asymptotically AdS,  $\Omega(r/L_i) \equiv r/\ell$ ,  $\ell$  being the AdS scale.

possible connection in the  $\mathcal{K} = 0$  case. Further, in the *special case* of a black brane solution in Einstein-Gauss-Bonnet theory, we have  $\kappa/r_H = (D-1)/2\ell^2$ , in which case the result above matches exactly with the one obtained in the context of AdS-CFT (see [62] and references within). Using the expression for  $\eta/s$ , it would be interesting to check whether it obeys or violates the KSS bound [66],  $\eta/s \geq 1/(4\pi)$ , by explicitly computing it for black hole solutions known for Lanczos-Lovelock gravity. For  $\mathcal{K} = 0$  case, the KSS bound is found to be violated in the AdS-CFT context in Einstein-Gauss Bonnet gravity [67, 62] and also in a general Lanczos-Lovelock theory of gravity [63, 68].

We can now summarize the broad implications of the results discussed in this (and the previous) section.

1. The connection between various transport coefficients and the thermodynamic properties of the horizon, such as temperature, entropy and quasi-local energy, becomes immediately obvious in our analysis, and it would be interesting if such a connection holds for horizons in more general class of gravity theories.
2. The calculation highlights the connection between Wald entropy  $S_{\text{Wald}}$  of the horizon and the pressure  $p_s$  of the horizon membrane. This relation is encoded in the form of an equation of state,

$$\frac{p_s A}{T_{\text{loc}}} \stackrel{N \rightarrow 0}{=} \sum_{m=1}^{[D-1]/2} \left( \frac{D-2m}{D-2} \right) S_{\text{Wald}}^{(m)} \quad (156)$$

The algebraic steps leading to the above result are precisely the same as those leading to the evaluation of entropy of a self-gravitating configuration of densely packed shells on the verge of becoming a black hole, discussed in the last section.

3. The relation between bulk and shear viscosities for Lanczos-Lovelock order  $m$ ,  $\zeta^{(m)}/\eta^{(m)} = -2(D-3)/(D-2)$  is also noteworthy. This ratio is *independent* of  $m$ , which immediately implies

$$\frac{\zeta}{\eta} = -2 \frac{D-3}{D-2} \quad (157)$$

for an action which is a *sum* of Lanczos-Lovelock terms. This corroborates the results already known for Einstein gravity, and recently also established for Einstein-Gauss-Bonnet case [62]. It would be worth investigating whether there is a deeper reason for the robustness of this ratio.

### 3.4 Second law for black hole entropy in Lanczos-Lovelock models

Having analyzed the Noether charge entropy in Lanczos-Lovelock models from various different perspectives in the previous sections, we now discuss the issue of whether the change in this entropy can be shown to be positive definite; that is, whether the second law can be shown to hold during a physical process for black hole entropy in Lanczos-Lovelock theory.

In general relativity, the “area theorem” implies that area of a black hole cannot decrease in any physical process, so long as the matter stress tensor satisfies the null energy condition [69]. However, the equilibrium state version of first law for black holes, which has been

established by Wald and collaborators, says nothing in general as to whether the Wald entropy always increases under physical processes in arbitrary higher curvature theories. In fact, black hole entropy is no longer proportional to area in arbitrary higher curvature theories in general, and Lanczos-Lovelock models in particular, and hence the question of whether the second law is valid in such cases is important and not yet solved (except for  $f(R)$  models, see [70]). The reference [70] also gives an argument for second law for arbitrary diffeomorphism invariant theories, but the argument relies on (a rather strong) assumption that stationary comparison version of the first law implies the physical process version for quasi-stationary processes.

We shall now sketch a recent proof of the validity of the second law for quasi-stationary physical processes involving black holes in Lanczos-Lovelock gravity [71]. Such a result was first proved for the Einstein-Gauss-Bonnet theory in [72], but since the result in [72] is a special case of the one proved in [71], we shall essentially outline the steps given in the latter reference. The main result of the analysis is that, Wald entropy of stationary black holes in Lanczos-Lovelock theories increases monotonically for physical processes which are quasi-stationary, at the end of which the black hole settles down to a quasi-stationary state. One of the reasons why this result is important is that it reinforces the physical fact that Wald entropy indeed behaves like ordinary thermodynamic entropy.

We shall concentrate on stationary, non-extremal, Killing horizons. In  $D$  spacetime dimensions, the event horizon is a null hypersurface  $\mathcal{H}$  with tangent vectors  $\mathbf{k} = \partial_\lambda$ , where  $\lambda$  is an affine parameter and  $\mathbf{k}$  satisfies geodesic equation. The constant  $\lambda$  slices foliate the horizon and are spacelike. An arbitrary point  $p$  on such slices can be assigned coordinates  $\{\lambda, x^A\}$  where  $x^A$ , ( $A = 2, \dots, D$ ) are coordinates of the point on  $\lambda = 0$  slice connected to  $p$  by a horizon generator. To complete the basis, one needs to introduce a second null vector  $\mathbf{l}$  such that  $\mathbf{l} \cdot \mathbf{k} = -1$ , thereby getting  $\{k^a, l^a, e_A^a\}$ . Finally, we would require the induced metric on any slice,  $\gamma_{ab} = g_{ab} + 2k_{(a}l_{b)}$ , with  $k^a\gamma_{ab} = 0 = l^a\gamma_{ab}$ . The change in induced metric from one slice to another is given by [73],

$$\mathcal{L}_k \gamma_{ab} = 2 \left( \sigma_{ab} + \frac{\theta}{(D-2)} \gamma_{ab} \right), \quad (158)$$

where  $\sigma_{ab}$  and  $\theta$  are shear and expansion, respectively, of the horizon. We will further assume that the event horizon is also a Killing horizon. This fact is true for general relativity due to strong rigidity theorem [74]. However, no such proof seems to exist for the Lanczos-Lovelock models to the best of our knowledge. With this assumption, the horizon is then generated by orbits of a Killing field  $\boldsymbol{\xi} = \partial_v$  which becomes null on the horizon, and surface gravity  $\kappa$  is defined in the standard manner as  $\xi^a \nabla_a \xi^b = \kappa \xi^b$ . For stationary spacetimes with a Killing horizon,  $\sigma_{ab} = 0; \theta = 0$ . Using evolution equation for  $\theta$  (Raychaudhuri equation) and shear, we obtain [73],

$$R_{ab} k^a k^b = \xi^a \gamma_i^b \gamma_j^c \gamma_k^d R_{abcd} = k^a k^c \gamma_m^b \gamma_n^d R_{abcd} = 0. \quad (159)$$

Now consider a physical process wherein a weak stress-energy tensor perturbs a stationary black hole for certain duration, and after which the black hole settles to a stationary state in asymptotic future. This last point implies that  $\boldsymbol{\xi}$  is an exact Killing vector in asymptotic future. Further, the flow of stress-energy is assumed to be slow, so that changes in dynamical fields are first order in some suitable (bookkeeping) parameter  $\epsilon$  and viscous effects can be neglected. To be more specific,  $\theta \sim \mathcal{O}(\epsilon), \sigma_{ab} \sim \mathcal{O}(\epsilon)$ .

In general relativity, a concrete example of such a physical process is slow accretion of matter by a black hole of mass  $M$  for a finite time, after which the black hole settles down to a stationary state. Linearized Raychaudhuri equation then gives,

$$\frac{d\theta}{d\lambda} \approx -R_{ab}k^ak^b = -8\pi T_{ab}k^ak^b, \quad (160)$$

where Einstein's equation have been used to obtain the second equality. If  $T_{ab}$  satisfies null energy condition  $T_{ab}k^ak^b \geq 0$ , the rate of change of  $\theta$  is negative on any slice prior to asymptotic future. Since  $\theta = 0$  in the future, the generators must have positive expansion during the accretion process. This suffices to prove the monotonic increase of area in the physical process. It is evident that the result depends crucially on field equations, and hence monotonicity of horizon area is only valid in case of general relativity.

For Lanczos-Lovelock gravity, the starting point, of course, is the Wald entropy of a stationary Killing horizon [42, 75, 43],

$$S = \frac{1}{4} \int \rho \sqrt{\gamma} dA \quad (161)$$

where the entropy density (after separating out the  $m = 1$  term) is given by (see Sec. 3.1):

$$\rho = \left( 1 + \sum_{m=2}^{[D-1]/2} 16\pi m \alpha_m L_{(m-1)}^{(D-2)} \right). \quad (162)$$

The aim is to prove that this entropy always increases when a black hole is perturbed by a weak matter stress energy tensor of  $\mathcal{O}(\epsilon)$  obeying null energy condition. To study the entropy change

$$\Delta S = \frac{1}{4} \int_{\mathcal{H}} \left( \frac{d\rho}{d\lambda} + \theta \rho \right) d\lambda \sqrt{\gamma} dA. \quad (163)$$

it is useful to define a quantity  $\Theta$  as,

$$\Theta = \left( \frac{d\rho}{d\lambda} + \theta \rho \right). \quad (164)$$

In general relativity,  $\Theta$  is the same as the expansion parameter  $\theta$  of null generators, but for a general Lanczos-Lovelock model,  $\Theta$  is the rate of change of the entropy associated with a infinitesimal portion of horizon ([70] discusses similar construction in  $f(R)$  gravity). Our aim is to prove that, given our assumptions,  $\Theta$  is positive on any slice in a physical process.

To proceed further, we use the fact already noted in Sec. 4.1.2, that the change of the  $(D-2)$ -dimensional scalar  $L_{(m-1)}^{(D-2)}$  can be thought of due to the change in the intrinsic metric, and this variation of  $L_{(m-1)}^{(D-2)}$  simply gives the corresponding equations of motion of  $(m-1)$ -th order Lanczos-Lovelock term in  $(D-2)$  dimensions. So, for a general Lanczos-Lovelock

gravity, we can write

$$\begin{aligned}\frac{d\rho}{d\lambda} &= \sum_{m=2}^{[D-1]/2} 16\pi m \alpha_m k^a \nabla_a (L_{(m-1)}^{(D-2)}) \\ &= - \sum_{m=2}^{[D-1]/2} 16\pi m \alpha_m {}^{(D-2)}\mathcal{R}_{(m-1)}^{ab} \mathcal{L}_k \gamma_{ab},\end{aligned}\quad (165)$$

where the surface term has been ignored since horizon sections are assumed to be compact surfaces without boundaries. The tensor  ${}^{(D-2)}\mathcal{R}_{ab}$ , which is the generalization of Ricci tensor for  $(m-1)$ -Lanczos-Lovelock term, is given by

$${}^{(D-2)}\mathcal{R}_{b(m-1)}^a = \frac{1}{16\pi} \frac{(m-1)}{2^m} \delta_{bc_1 d_1 \dots c_m d_m}^{a_1 b_1 \dots a_m b_m} {}^{(D-2)}R_{a_1 b_1}^{ad_1} \dots {}^{(D-2)}R_{a_m b_m}^{cd_m} \quad (166)$$

Using Eq. (158), one now obtains,

$$\Theta = \theta + 16\pi \sum_{m=2}^{[D-1]/2} \alpha_m \left[ -2 \left( \frac{{}^{(D-2)}\mathcal{R}_{(m-1)} \theta}{(D-2)} + {}^{(D-2)}\mathcal{R}_{(m-1)}^{ab} \sigma_{ab} \right) + \theta L_{(m-1)}^{(D-2)} \right]. \quad (167)$$

To study the rate of change of  $\Theta$  along the congruence, we use the evolution equations for expansion (the Raychaudhuri equation) and shear [73]. Since we work to first order in perturbation over a background stationary spacetime, we can always extract the part linear in perturbation from a product of quantities  $XY$  as

$$XY \approx X^{(B)} Y^{(P)} + X^{(P)} Y^{(B)} \quad (168)$$

where  $X^{(B)}$  is the value of  $X$  evaluated on the stationary background, and  $X^{(P)}$  is the linear part of its perturbation. Note that, for stationary background, Raychaudhuri equation demands  $R_{ab}^{(B)} k^a k^b = 0$ . Since  $T_{ab}^{(B)} k^a k^b = 0$ , this implies  $E_{(m)ab}^{(B)} k^a k^b = 0$ . We shall further simplify the calculation by using diffeomorphism freedom to make the null geodesic generators of the event horizon of the perturbed black hole coincide with the null geodesic generators of the stationary black hole [76].

Using the evolution equation of  $\theta$  and  $\sigma_{ab}$  to linear order:  $d\theta/d\lambda \approx -R_{ab}^{(P)} k^a k^b$  and  $d\sigma_{ab}/d\lambda \approx C_{acdb}^{(P)} k^c k^d$ , and Eq. (159) for the background, the evolution of  $\Theta$  to linear order in perturbation can be written as

$$\frac{d\Theta}{d\lambda} = -8\pi T_{ab} k^a k^b + \mathcal{D}_{ab} k^a k^b, \quad (169)$$

where

$$\mathcal{D}_{ab} k^a k^b = \sum_{m=2}^{[D-1]/2} 16\pi \alpha_m \left[ E_{(m)ab}^{(P)} k^a k^b + 2m {}^{(D-2)}E_{(m-1)}^{(B)ab} R_{acbd}^{(P)} k^c k^d \right]. \quad (170)$$

and we have used the expression of perturbed Weyl tensor in terms of curvature and Ricci tensors, along with the relation  ${}^{(D-2)}E_{(m)ab} = {}^{(D-2)}\mathcal{R}_{(m)ab} - (1/2)\gamma_{ab} {}^{(D-2)}L_{(m)}$ .



The next, and the most crucial, step is to prove that the first order part of  $\mathcal{D}_{ab}k^ak^b$  vanishes identically. We omit the proof of this, which can be found in [71]. It is, however, worth mentioning that the proof depends crucially on various algebraic properties of Lanczos-Lovelock models that have been discussed in earlier parts of this review. Whether there exists another way of proving the result which is adaptable to gravity theories more general than Lanczos-Lovelock remains an open issue.

Once the first order part of  $\mathcal{D}_{ab}k^ak^b$  is shown to vanish identically, we finally arrive at,

$$\frac{d\Theta}{d\lambda} = -8\pi T_{ab}k^ak^b + \mathcal{O}(\epsilon^2). \quad (171)$$

The above equation shows that if null energy condition holds, the rate of change of  $\Theta$  is always negative during a slow classical dynamical process which perturbs the black hole to a new stationary state. Since, the final state by assumption is stationary, both  $\theta$  and  $\sigma$ , and as a consequence  $\Theta$ , vanishes in the asymptotic future. We can use therefore use the same argument as in the case of general relativity to conclude that  $\Theta$  must be positive at every slice during the process, thereby proving that the horizon entropy of black holes in Lanczos-Lovelock gravity is indeed a monotonically increasing function during any quasi-stationary physical process, i.e.

$$\frac{dS}{d\lambda} \geq 0. \quad (172)$$

Further comments on the implications of this result, and suggestion for further investigations, can be found in [71].

### 3.5 Entropy from horizon Virasoro algebra

There is now considerable evidence to show that features like entropy, temperature etc—which were originally thought of as properties of black holes [37, 38]—arises in a much more general context. Several investigations suggest that there exist certain level of universality in the thermodynamic properties of null surfaces. Hence approaches which compute the entropy density of a null surface in a generic, universal, manner might provide us with deeper insight into the quantum structure of spacetime. This aspect was emphasized, among others, by Carlip who has developed such an approach to compute the entropy using horizon Virasoro algebra.

The basic idea is based on an approach first presented by Brown and Henneaux [77], which was originally used in the context of asymptotically AdS space-time in  $(1+2)$  dimensional gravity. They found that the Fourier modes  $Q_j$  of the charges corresponding to the asymptotic diffeomorphism symmetry generators obey a Virasoro algebra with central extension:

$$i[Q_m, Q_n] = (m-n)Q_{m+n} + \frac{C}{12}m^3\delta_{m+n,0}, \quad (173)$$

where  $C$  is known as the central charge. The work by Strominger [78] and others showed that if one uses the above central charge in the Cardy formula [79, 80], the resulting entropy comes out to be the Bekenstein-Hawking entropy for the  $(1+2)$  dimensional black hole. This was further developed by Carlip [81, 82] using the diffeomorphism symmetry generators *near the horizon* to obtain the black hole entropy. In this approach, one begins with the

diffeomorphism generators  $\xi_n^a(x)$  which preserve a set of boundary conditions near the horizon. The Fourier modes of these generators obey one sub-algebra isomorphic to  $\text{Diff } S^1$  given by

$$i\{\xi_m, \xi_n\}^a = (m - n)\xi_{m+n}^a \quad (174)$$

where  $\{, \}$  is the Lie bracket. One can then construct the Fourier modes  $Q_n$  of the charges corresponding to each  $\xi_n^a$ , either by Hamiltonian [83] or a covariant Lagrangian formalism [84, 41, 85, 86, 87, 88] and evaluate the Lie brackets among them. A comparison between this algebra and Eq. (173) allows us to identify the central charge. Finally, one finds the zero mode eigenvalue  $Q_0$  and computes the entropy of the black hole using Cardy formula.

Several related approaches have been developed using these ideas with the hope that diffeomorphism symmetry generators may shed some light towards the microscopic degrees of freedom responsible for entropy of the horizon [89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103]. All these approaches developed in the literature have the following ingredients:

- The Noether current used in the approaches are defined on-shell, usually by ignoring a term which vanishes when equations of motion are used.
- The calculation of Lie brackets is complicated and different approaches lead to slightly different results and it is not clear how to interpret these differences in the calculations.
- To obtain the correct result, one has to further impose specific boundary conditions on the horizon in order to set certain terms to zero. Again, the physical meaning of these boundary conditions is often not clear.
- Most of the analysis (except the one in [99]) is confined to general relativity and it is not clear how to generalize the results for a wider class of theories.

We will now describe an approach to horizon entropy based on the existence of a Virasoro algebra and its central charge as developed in Ref. [104] in which it was shown that there is a relatively simple way of obtaining the central charge and the horizon entropy using the off-shell Noether current in any generally covariant theory of gravity, and in particular for the Lanczos-Lovelock model. The method does not require any boundary conditions to make unwanted terms to vanish. The essential idea is to use diffeomorphism invariance of the Lagrangian under  $x^i \rightarrow x^i + \xi_1^i$  to define the Noether current  $J_a[\xi_1]$  and then use *its* variation  $\delta_{\xi_2} J_a[\xi_1]$  under a second diffeomorphism  $x^i \rightarrow x^i + \xi_2^i$  to define the Lie bracket structure. This can be done without using the explicit form of Noether current, and the calculation of the resulting Lie bracket is algebraically quite simple and leads to the standard results. One can then identify the resulting Virasoro algebra, the central charge and the zero-mode eigenvalue; plugging these into the Cardy formula, one obtains entropy density of a null surface which turns out to be the same as Wald entropy for the Lanczos-Lovelock models.

The key new features of this approach are as follows:

- The current is now *defined and conserved* off-shell i.e. we do not use the equations of motion in its definition and in its conservation. A derivation of Noether current and its conservation for *on-shell* condition is given in [105]. In this case a specific boundary condition is required to obtain the necessary results. But the approach in Ref. [104] deals with off-shell situation and does not require any ad-hoc boundary condition. This is in contrast to — and an improvement on — the previous approaches which use equations of motion extensively.

- The definition of the bracket among the charges is completely general in the sense that one does not need the explicit expression for the Noether current to obtain its structure.
- This approach also provides a derivation of the central term (and zero mode energy) from the surface term contribution  $\delta v^a$  in the Noether current. This is new and does not seem to have been noticed in the previous works.
- In the derivation of bracket we do not use the equations of motion or specific boundary conditions (like Dirichlet or the Neumann boundary condition) to make non-covariant terms vanish; therefore this approach is general enough to include any covariant theory of gravity. (This feature was again missing in the earlier literature which used the equations of motion for the evaluation of the relevant brackets as an on-shell construct.) As a direct consequence of the above, the central term obtained is also an off-shell construct.

Consider any generally covariant Lagrangian, for which the conserved Noether current  $J^a$  (with a corresponding current density  $P^a$ ) that can be expressed as the covariant derivative of an anti-symmetric Noether potential  $J^{ab}$ . These satisfy the standard conservation laws which are valid off-shell:

$$J^a \equiv \nabla_b J^{ab}, \quad P^a \equiv \sqrt{g} J^a; \quad \nabla_a J^a = 0, \quad \partial_a P^a = 0. \quad (175)$$

Let us now consider the variation of the current density itself for an arbitrary diffeomorphism  $x^a \rightarrow x^a + \xi^a$ . and use it to obtain the variation of the Noether charge by straight forward calculation:

$$\delta_\xi Q \equiv \int d\Sigma_a \delta_\xi P^a = \int d\Sigma_{ab} \sqrt{h} \xi^b J^a, \quad (176)$$

where  $h$  is the determinant of the induced metric of  $(d-2)$ -dimensional boundary in a  $d$ -dimensional spacetime. We next define a (Lie) bracket structure for the charges which leads to the usual Virasoro algebra with the central extension for a general Lagrangian of the kind  $L = L(g_{ab}, R_{abcd})$ . The relevant bracket among the charges is defined as:

$$[Q_1, Q_2] \equiv (\delta_{\xi_1} Q[\xi_2] - \delta_{\xi_2} Q[\xi_1]) \quad (177)$$

Then using Eq. (176), we obtain:

$$[Q_1, Q_2] := \int \sqrt{h} d\Sigma_{ab} [\xi_2^a J^b[\xi_1] - \xi_1^a J^b[\xi_2]] \equiv \int \sqrt{h} d\Sigma_{ab} [\xi_2^a J_1^b - \xi_1^a J_2^b] \quad (178)$$

where we use the notation  $J_1^b = J^b[\xi_1]$  etc. We note that: (a) This definition is quite general and has not used any field equations. (b) The derivation of Eq. (178) is simple and straightforward. (c) If we use the form  $J^a = \nabla_b (\nabla^a \xi^b - \nabla^b \xi^a)$  for general relativity one can easily obtain the result obtained earlier in the literature, like e.g., in Ref. [82]. We will hereafter concentrate on Lanczos-Lovelock gravity and evaluate Eq. (178) on the  $(d-2)$ -dimensional null surface which is a Killing horizon.

To evaluate Eq. (178) over the Killing horizon, we will follow the stretched horizon approach of Carlip [82]. Let us first mention some of the key results needed for this computation.

The location of the horizon is defined by the vanishing of the norm of a timelike (approximate) Killing vector  $\chi^a$ . Near the horizon, one can define a vector  $\rho^a$ , orthogonal to the orbits of the Killing vector  $\chi^a$ , by the following relation

$$\nabla_a \chi^2 = -2\kappa \rho_a , \quad (179)$$

where  $\kappa$  is the surface gravity at the horizon, with  $\chi^a \rho_a = 0$ . Consider a class of diffeomorphism generators given by:

$$\xi^a = T\chi^a + R\rho^a . \quad (180)$$

where  $T$  and  $R$  are scalar functions chosen such that the generators obey the (near-horizon) condition  $[(\chi^a \chi^b)/\chi^2] \delta_\xi g_{ab} \rightarrow 0$  which preserves the horizon structure. This condition leads to a relation among  $R$  and  $T$  given by:

$$R = \left( \frac{\chi^2}{\rho^2 \kappa} \right) \chi^a \nabla_a T \equiv \frac{\chi^2}{\kappa \rho^2} D T , \quad (181)$$

where  $D \equiv \chi^a \nabla_a$ . We can now compute the Noether charge:

$$\begin{aligned} Q[\xi] &= -\frac{1}{32\pi G} \int \sqrt{h} d^{d-2} X \mu_{ab} \frac{\rho}{|\chi|} P^{abcd} \mu_{cd} \left[ 2\kappa T - \frac{1}{\kappa} D^2 T \right] \\ &= -\frac{1}{32\pi G} \int \sqrt{h} d^{d-2} X P^{abcd} \mu_{ab} \mu_{cd} \left[ 2\kappa T - \frac{1}{\kappa} D^2 T \right] \end{aligned} \quad (182)$$

and the central term of the Virasoro algebra defined by the relation,

$$K[\xi_1, \xi_2] = [Q_1, Q_2] - Q[\{\xi_1, \xi_2\}] , \quad (183)$$

where  $[Q_1, Q_2]$  is given by Eq. (177). We find that the central term is given by

$$K[\xi_1, \xi_2] = -\frac{1}{32\pi G} \int \sqrt{h} d^{d-2} X P^{abcd} \mu_{ab} \mu_{cd} \frac{1}{\kappa} \left[ D T_1 D^2 T_2 - D T_2 D^2 T_1 \right] . \quad (184)$$

This was obtained earlier in Ref.[99] by a more complicated procedure and using the on-shell expressions. We have obtained it without using the field equations or any special boundary conditions thereby demonstrating the generality of the result.

Using a suitably defined Fourier decomposition of the  $T_1$  and  $T_2$  in Eq. (184) we can find the central charge and zero mode eigenvalue. We use a Fourier decomposition of  $T_1$  and  $T_2$  given by:

$$T_1 = \sum_m A_m T_m; \quad T_2 = \sum_n B_n T_n , \quad (185)$$

with  $A_n^* = A_{-n}$ ,  $B_m^* = B_{-m}$  and

$$T_m = \frac{1}{\alpha} \exp[i m (\alpha t + g(x) + p \cdot x_\perp)] \quad (186)$$

where  $\alpha$  is a constant,  $g(x)$  is a function that is regular at the Killing horizon,  $p$  is an integer and  $x_\perp$  are the  $(d-2)$  tangential coordinates. Here the  $t-x$  plane defines the null surface.

Obviously, the result will depend on the choice made for  $\alpha$  and we need to fix this to get a unique value for entropy. A natural choice, arising from the fact that near-horizon Rindler geometry exhibits periodicity in imaginary time with period  $2\pi/\kappa$ , is

$$\alpha = \kappa \quad (187)$$

Substituting Eq. (186) in Eq. (184) and defining a quantity

$$\hat{\mathcal{A}} = -\frac{1}{2} \int \sqrt{h} d^{d-2} X P^{abcd} \mu_{ab} \mu_{cd} , \quad (188)$$

which is proportional to the Wald entropy, we obtain,

$$K[\xi_m, \xi_n] := -im^3 \frac{\hat{\mathcal{A}}}{8\pi G} \delta_{n+m,0} . \quad (189)$$

Similarly, using the Fourier decomposition,  $Q[\xi] = \sum_m A_m Q[\xi_m]$  in Eq. (182), we obtain,

$$Q[\xi_m] = \frac{\hat{\mathcal{A}}}{8\pi G} \delta_{m,0} . \quad (190)$$

and the algebra of charges turns out to be:

$$i[Q_m, Q_n] = (m-n)Q[\xi_{m+n}] + m^3 \frac{\hat{\mathcal{A}}}{8\pi G} \delta_{n+m,0} . \quad (191)$$

This is the standard form of the Virasoro algebra in Eq. (173) with  $Q[\xi_{m+n}] \equiv Q_{m+n}$ . We can identify the central charge and the zero mode eigenvalue as:

$$\frac{C}{12} = \frac{\hat{\mathcal{A}}}{8\pi G}; \quad Q[\xi_0] = \frac{\hat{\mathcal{A}}}{8\pi G} . \quad (192)$$

The standard Cardy formula for the entropy is given by [79, 80],

$$S = 2\pi \sqrt{\frac{C\Delta}{6}} ; \quad \Delta \equiv Q_0 - \frac{C}{24} \quad (193)$$

which leads to

$$S = \frac{\hat{\mathcal{A}}}{4G} \quad (194)$$

This exactly matches with the Wald entropy of the theory.

The conceptual basis of this approach was concretely demonstrated (in the case of general relativity) from another point of view [106]. One can show that the Noether current that arises from the *surface term* of the action in general relativity, for diffeomorphisms which preserve the horizon structure, also possess a Virasoro algebra with similar properties. More precisely ref. [106] considers a class of diffeomorphisms which leave a particular form of the Rindler metric invariant and introduces the Noether current associated with the *surface term* in the

action functional for gravity (in contrast to previous approaches in the literature in which either the bulk term or the Einstein-Hilbert action was used to define the Noether current). Given the Noether current and diffeomorphism generators, there is a natural Virasoro algebra which can be associated with the horizon. The central extension of this Virasoro algebra again leads — via Cardy formula — to the correct entropy density of the horizon. Just to provide a contrast with the previous analysis, we will mention couple of technical points. In the earlier work, to obtain the correct entropy one had to either shift the zero mode energy (as done in, e.g., [82]) or choose a parameter contained in the Fourier modes of  $T$  (for instance  $\alpha$ ) as the surface gravity (as done in, e.g., [103]) or both [104] depending on the action of the theory we have started with. *Here, interestingly, one did not require any such ad hoc prescription.* This is because the parameter  $\alpha$  in the expression of  $T_m$  does *not* appear in the final expression of entropy in this approach.

Using (a) the Noether current corresponding to the York-Gibbons-Hawking surface term and (b) choosing the vector fields by demanding invariance of the Rindler form of the metric, have not been attempted before and are new features of the work in Ref. [106]. This is important because it further strengthens the peculiar feature present in all gravitational actions, viz., that the same information is encoded in both bulk and boundary terms of the action.

From a conceptual point of view, the analysis in Ref. [106] provides a nice, simple, physical picture of the connection between horizon entropy and the degrees of freedom which contribute to it. In conventional physics, we are accustomed to thinking of degrees of freedom of a system as absolute and independent of observer or the coordinate system used by the observer. In such a picture, the entropy — which is related to the logarithm of the degrees of freedom — will also be absolute and independent of the observer. We however know that horizon entropy, horizon temperature etc. must be treated as observer dependent notions because a freely falling observer through a black hole horizon and a static observer outside the black hole will attribute different thermodynamic properties to the horizon. *It follows that any microscopic degrees of freedom which leads to horizon entropy must also be necessarily observer and coordinate dependent.* So the question “what are the degrees of freedom responsible for black hole entropy?” has no observer independent answer. We find that the notion of residual gauge symmetries which are perceived as real degrees of freedom under a restricted class of diffeomorphism, allows us to quantify this notion. An observer who perceives a horizon works with a theory having less diffeomorphism symmetry if she wants to retain the structure of the metric near the horizon. This necessarily upgrades some of the original gauge degrees of freedom to effectively true degrees of freedom as far as this particular observer is concerned. As a result, this particular class of observers will attribute an entropy to the horizon in an observer dependent manner.

The result of [106] has already been generalised to  $m = 2$  and  $m = 3$  Lanczos-Lovelock terms in [107], using the corresponding surface terms for these lagrangians given in Sec. 2.6, and we expect it to hold for all Lanczos-Lovelock terms in a straightforward manner [108].

## 4 Lanczos-Lovelock models and the emergent gravity paradigm

Recent results indicate that gravity could be an emergent phenomenon, like fluid mechanics or elasticity, with the field equations governing gravitational dynamics having the same status as, say, the equations of fluid mechanics. This alternative perspective for classical gravity appears to be conceptually more satisfying than the usual perspective in certain aspects. (See for a review, Refs. [3, 4], [109].) and several features of Lanczos-Lovelock models strengthen this paradigm. In this section we shall describe these features.

The back drop in which such an approach arises is the following: it is reasonable to assume that there are certain pre-geometric variables underlying the spacetime structure and that there exists a microscopic theory describing their dynamics. If we think of continuum spacetime as analogous to a continuum description of a fluid, the microscopic — quantum gravitational — description is analogous to statistical mechanics of the molecules of the fluid. (That is, the pre-geometric variables will act as the “atoms of spacetime”.) The exact quantum theory of gravity, when it is available, will allow us to construct, in a coarse-grained long wavelength limit, a smooth spacetime and the effective degrees of freedom (like the metric tensor) in terms of pre-geometric variables. This procedure is analogous to obtaining variables like pressure, temperature etc. for a fluid in terms of microscopic variables. The dynamical equations of the microscopic theory will also lead to some effective description for the emergent degrees of freedom. For example, given the microscopic description of molecules moving randomly inside a container and colliding with the walls, one can obtain the ideal gas law  $PV = Nk_B T$  governing the macroscopic (“emergent”) variables like pressure, temperature etc. Similarly, once we have the correct theory of quantum gravity (in terms of pre-geometric variables), we should be able to obtain the field equations governing the geometry by a suitable approximate description.

Such an approach, however, works only when we know the underlying microscopic theory. How do we proceed when we do *not* know the underlying theory, which is the current situation as regards the spacetime? In the case of normal matter, this was indeed the situation couple of centuries back and physicists could successfully describe the behaviour of matter using thermodynamic concepts, even before we understood the molecular structure of matter. The emergent paradigm for gravity can be viewed as a similar attempt to describe gravity in a thermodynamic language, given the fact that we do not know the exact, microscopic description of spacetime.

In the case of normal matter, the evidence for the existence of microscopic degrees of freedom is provided by the elementary observation that matter could be heated up. Boltzmann emphasized the fact that matter could store and transfer energy in the form of heat *only because* it had microscopic degrees of freedom. If a fluid admits continuum description all the way and has no internal degrees of freedom, it cannot store heat or have a notion of temperature. Thus the existence of thermal phenomena provides strong internal evidence that continuum fluid mechanics misses some essential aspect of physics. This is clearly seen, in the case of an ideal gas, if we write

$$Nk_B = \frac{PV}{T} \tag{195}$$

The quantities on the right hand side of Eq. (195) are well defined and meaningful in ther-

modynamics but the  $N$  in the left hand side has no physical meaning in thermodynamics. It is proportional to the number of microscopic degrees of freedom of the gas and is irrelevant in the strictly continuum limit. With hindsight, we now know that the existence of the gas constant  $R = Nk_B$  for a mole of gas has a microscopic interpretation related to the internal degrees of freedom. This is a situation in which we could have guessed (as Boltzmann did) the existence of microscopic degrees of freedom — and even counted it, as Avogadro number for a mole of gas — from the existence of thermal phenomena in normal matter.

In a similar vein, we can ask whether the continuum description of spacetime dynamics *exhibits evidence* for a more fundamental level of microscopic description in terms of “atoms of spacetime”. It would, of course, be impossible to describe the statistical mechanics of such microscopic degrees of freedom knowing only the continuum description of spacetime, just as one could not understand the constituents of matter in the days of Boltzmann. But it must be possible to provide internal evidence strongly suggesting the *existence of* such atoms of spacetime just by probing the continuum theories of gravity.

The first task of the emergent paradigm will be to come up with such ‘internal evidence’ by closely examining classical theories of gravity. We will discuss several peculiar features in the structure of classical gravitational theories which have no deeper explanation within the standard framework and have to be accepted as just algebraic accidents. These peculiar features provide us with hints about the underlying microscopic theory. In that sense, they are conceptually similar to the equality of inertial and gravitational mass (which could have been thought of as an algebraic accident in Newtonian gravity but does find a deeper explanation when gravity is described as spacetime geometry) or the fact that matter can be heated up (which defied a fundamental explanation until Boltzmann postulated the existence of microscopic degrees of freedom). These features strongly suggest interpreting classical gravity as an emergent phenomenon with its field equations having the same status as equations of, say, fluid mechanics. Careful analysis of classical gravity (with a single quantum mechanical input, viz., the Davies-Unruh temperature [39] of local Rindler horizon) leads us to this conclusion.

In the case of normal matter, a more formal link between microscopic and macroscopic description is established by specifying certain thermodynamic potentials like entropy, free-energy etc. as suitable functions of emergent variables like pressure, temperature etc. The functional form of these potentials can, in principle, be derived from the microscopic theory; but, when we do not know the microscopic theory, they are postulated phenomenologically from the known behaviour of the macroscopic systems. Similarly, we would expect such a thermodynamic approach to work in this case of spacetime as well. It should be possible to express, say, an entropy or free energy density for spacetime in terms of suitable variables, the extremum of which should lead to a consistency condition on the background spacetime — which will act as the equation of motion. Such a consistency condition arises even in normal thermodynamics, though it is not often described as one such. For a gas of  $N$  molecules in a volume  $V$ , we can express the kinetic energy and momentum transfer through collisions to the walls of the container per unit area and time, entirely in terms of the microscopic variables. Coarse graining these quantities we obtain the macroscopic variables  $T$  and  $P$ .



The equations of microscopic physics now demand the *consistency condition*

$$\frac{\left( \begin{array}{c} \text{mean momentum transfer to walls} \\ \text{per unit time per unit area} \end{array} \right)}{\left( \begin{array}{c} \text{mean kinetic energy} \end{array} \right)} = \frac{N}{V} = \frac{P}{T} \quad (196)$$

between the coarse grained variables. We will show in Sec. 4.2 that such an approach is indeed possible to obtain the field equations of gravity.

It is also true that, in the case of normal matter, the laws of thermodynamics hold universally and do not depend on the kind of matter (ideal gas, liquid crystal, metal, ... ) one is studying. The information about the *specific* kind of matter which one is studying is provided by the *specific* functional form of the thermodynamic potential like e.g., the free energy  $F = F(T, V)$ . Similarly, the thermodynamic framework can describe a wide class of possible gravitational field equations for the effective degrees of freedom. Which of these actually describe nature depends on the specific functional form of the thermodynamic potential, say, the entropy density of spacetime. It turns out that, for the Lanczos-Lovelock models which we will concentrate on, this information is encoded in the tensor  $P_{cd}^{ab}$  (which we will call the *entropy tensor*) and the nature of the resulting theory depends on the dimension of spacetime. In particular, if  $D = 4$ , the thermodynamic paradigm selects Einstein's theory uniquely.

This approach can be thought of as a “top-down” view (in real space) from classical gravity to quantum gravity. Specific quantum gravitational models which approach the problem ‘bottom-up’ has to be consistent with the features of classical gravity described in the sequel. In particular, this approach makes precise the task of the microscopic quantum gravity model viz., that it should lead to a specific functional form for the entropy or free energy density of spacetime, just as microscopic statistical mechanics will lead to a specific entropy (or free energy) functional for a material system. In this sense, the emergent approach complements microscopic approaches based on models of quantum gravity. Such a perspective also implies that quantizing any classical gravitational field will be similar to quantizing, say the equations of fluid dynamics. Gravitons will have the same status as phonons in a solid. Neither will give us insights into the deeper microstructure — spacetime or atoms. (The alternative paradigm also has important implications for cosmology and, in fact, suggests linking cosmology with the emergence of space itself in a special manner. The details can be found in ref. [109]. )

## 4.1 Emergent aspects of the Lanczos-Lovelock equations of motion

The first ‘internal evidence’ suggesting that Lanczos-Lovelock models describe an emergent structure of spacetime arises from the remarkable fact that Lanczos-Lovelock field equations reduce to a thermodynamics identity  $TdS = dE + PdV$  on the horizon in any static spacetime. We will begin by describing this result. Obviously, it will depend crucially on the near horizon properties of certain tensors associated with Lanczos-Lovelock field equations which we will describe in Sec. 4.1.1. These will be shown to lead to the thermodynamics identity  $TdS = dE + PdV$  on the horizon, in Sec. 4.1.2.

### 4.1.1 Near-horizon symmetries of equation of motion

To investigate the near-horizon symmetries, it is convenient to use a setup which is best suited to describe geometry in the neighbourhood of a static horizon. We will describe the

static case (viz., the existence of a timelike Killing vector at least in the neighbourhood of the horizon); the more general case of stationary horizons will only be mentioned when we require it.

A coordinate system which facilitates taking the near horizon limit is given by the Gaussian normal coordinates, which, for static spacetimes, simplifies further and can be written as [110]

$$ds^2 = -N^2 dt^2 + dn^2 + \sigma_{AB} dy^A dy^B \quad (197)$$

where  $\sigma_{AB}(n, y^A)$  is the transverse metric, and the Killing horizon, generated by the timelike Killing vector field  $\xi = \partial_t$ , is approached as  $N^2 \rightarrow 0, n \rightarrow 0$ . We give below some geometric quantities associated with the above metric, which will be put to good use in subsequent analysis.

1. *Taylor expansion of metric components:* Demanding the finiteness of curvature invariants on the horizon leads to the following Taylor series expansion for the lapse  $N(n, y)$  and the transverse metric  $\sigma_{AB}$  [110]:

$$\begin{aligned} N(n, y) &= \kappa n \left[ 1 - \frac{1}{2} R_{\perp}(y; n=0) n^2 + O(n^3) \right] \\ \sigma_{AB} &= [\sigma_H(y)]_{AB} + \frac{1}{2} [\sigma_2(y)]_{AB} n^2 + O(n^3) \end{aligned} \quad (198)$$

where we have collectively called the transverse coordinates as  $y$ , and  $R_{\perp}$  is the Ricci scalar corresponding to  $y = \text{constant}$  surface.

2. *Domain of validity:* We shall be interested in a small region in the neighborhood of the spacelike  $(D-2)$ -surface; more precisely, we shall assume  $n \ll R_{\perp}^{-1/2}$ , and would be interested in the  $\kappa \rightarrow \infty$ . That is, we shall require the length scale set by  $\kappa$  to be the smallest of all length scales in the problem. Note that the coordinate  $n$  gives the normal distance from the horizon.
3. *The null geodesic congruences:* Consider the null vectors given by:

$$\begin{aligned} l_i &= (-1, +N^{-1}, 0, 0) \\ k_i &= (-1, -N^{-1}, 0, 0) \end{aligned} \quad (199)$$

Explicit calculation shows that  $\nabla_{\mathbf{l}} \mathbf{l}$  has only “ $y$ ”-components, given by:

$$\nabla_{\mathbf{l}} \mathbf{l} = - (2\kappa^{-2}) [\sigma^{AB} \partial_B R_{\perp}]_{n=0} \partial_A + O(n^2) \quad (200)$$

which go to zero in the limit  $\partial R_{\perp}/\kappa^3 \rightarrow 0$ ; therefore, in the limit we are interested in, it is easy to check that these vectors satisfy the geodesic equation in affinely parametrized form, that is;  $\nabla_{\mathbf{l}} \mathbf{l} = 0 = \nabla_{\mathbf{k}} \mathbf{k}$ .

The affine parameter  $\lambda$ , defined by  $\mathbf{l} \cdot \nabla \lambda = 1$  can be found by using the above form of  $N(n, y)$ ; to the leading order, we find that,  $\lambda \sim \lambda_H + (1/2)\kappa n^2$  where  $\lambda = \lambda_H$  is the

location of the horizon. Note that,  $N^2 \mathbf{l} \rightarrow \boldsymbol{\xi}|_H$ , which implies,  $2\kappa(\lambda - \lambda_H) \mathbf{l} \rightarrow \boldsymbol{\xi}|_H$ . In some part of the subsequent analysis, the differentials of various geometric quantities (such as entropy) defined on the horizon, which are directly involved in the statement of the first law of thermodynamics, would be interpreted as variations with respect to the affine parameter along the outgoing null geodesics, i.e.,  $\lambda$ . All of above comments hold (except for few sign changes) also for the ingoing null vector  $\mathbf{k}$ .

4. *Comoving observers and their acceleration:* Let us next look at the comoving observers in this spacetime,  $n = \text{const}$ ,  $y = \text{const}$ , with 4-velocity  $\mathbf{u} = N^{-1} \partial_t$ . It is easy to check that their acceleration  $\mathbf{a} = \nabla_{\mathbf{u}} \mathbf{u}$  is given by:

$$\mathbf{a} = \frac{1}{N^2} g^{\mu\nu} \partial_\nu N \partial_\mu \quad (201)$$

The components can be worked out explicitly by using Taylor series expansions in Eqs. (198); these are given by [with  $C = -(1/2)R_\perp(y; n = 0)$ ]

$$\begin{aligned} a_n &= a^n = n^{-1} + 2Cn + O(n^2) \\ a_\perp &= (\partial_\perp C)n^2 + O(n^4) \\ a^2 &= a_n^2 + \sigma^{AB} a_A a_B \\ &= n^{-2} + 4C + O(n^2) + O(n^4) \end{aligned}$$

so that

$$\frac{a^n}{a} = 1 + O(n^3) ; \quad \frac{a_\perp}{a} = O(n^3) \quad (202)$$

Therefore, we see that,

$$\hat{\mathbf{a}} = [1 + O(n^3)] \partial_n + O(n^3) \partial_A \quad (203)$$

This result indicates how the curvature  $R_\perp$  of the  $t - n$  part of spacetime, which is a plane spanned by  $\mathbf{u}$  and  $\mathbf{n}$ , will in general tend to produce an acceleration in the “ $y$ ” directions. We see that even in a curved spacetime, the unit vector along  $\mathbf{a}$  points along  $\mathbf{n}$  to an excellent approximation [ $O(n^3)$ ].

Getting back to the symmetries of equations of motion, we can now use the above near-horizon form of the metric to show that the Lanczos-Lovelock equations of motion satisfy the following near-horizon symmetries, generalizing the result first derived in [110]:

$$E_{\hat{b}}^{\hat{a}}|_H = \left[ \begin{array}{cc|c} E_\perp & 0 & 0 \\ 0 & E_\perp & 0 \\ \hline 0 & 0 & E_{\parallel \hat{B}}^{\hat{A}} \end{array} \right] \quad (204)$$

where  $E_\perp$  is given by

$$E_\perp = \frac{1}{16\pi} \frac{m}{2^m} [\sigma_2]_{CB} \sigma^{CA} \mathcal{E}_A^B - \frac{1}{2} L_m^{(D-2)} \quad (205)$$

$$\mathcal{E}_A^B = \delta_{AC_1 \dots D_{m-1}}^{BA_1 \dots B_{m-1}} {}^{(D-2)} R_{A_1 B_1}^{C_1 D_1} \dots {}^{(D-2)} R_{A_{m-1} B_{m-1}}^{C_{m-1} D_{m-1}} \quad (206)$$

In Sec. 4.1.2, we will use the above form when proving that Lanczos-Lovelock field equations reduce to  $TdS = dE + P dV$ . Here, we give a brief sketch of how the above, highly symmetric, form of Lanczos-Lovelock equations of motion arise near a static horizon as a result of the geometric structure of the Lanczos-Lovelock Lagrangian. More details can be found in the appendices of Ref. [111].

Let us first briefly outline the proof for Einstein tensor. We begin with the following expressions for the decomposition of the Riemann tensor for the metric in Eq. (197) (see, for example, Section 21.5 and Exercise 21.9 of [26]).

$$\begin{aligned} R^t_{\mu\nu\rho} &= 0 \\ R_{\mu t\nu t} &= NN_{|\mu\nu} \\ R^\mu_{\nu\rho\sigma} &= {}^{(3)}R^\mu_{\nu\rho\sigma} \end{aligned} \quad (207)$$

and

$$\begin{aligned} {}^{(3)}R_{ABCD} &= {}^{(2)}R_{ABCD} - (K_{AC}K_{BD} - K_{AD}K_{BC}) \\ {}^{(3)}R_{nBCD} &= K_{AC:B} - K_{AB:C} \end{aligned} \quad (208)$$

Here,  $|$  and  $:$  are the covariant derivatives compatible with the induced metric on  $\{t = \text{constant}\}$  and  $\{t = \text{constant}, n = \text{constant}\}$  surfaces respectively, and  $K_{AB} = -(1/2)\partial_n\sigma_{AB}$  is the extrinsic curvature of the  $\{t = \text{constant}, n = \text{constant}\}$  2-surface as embedded in the  $\{t = \text{constant}\}$  surface. These expressions lead to

$$\begin{aligned} R^A_{nB} &= \sigma^{AC} [\partial_n K_{CB} + (K^2)_{BC}] \\ R^A_{tB} &= \sigma^{AC} \left[ -\frac{N_{:CB} - K_{CB}\partial_n N}{N} \right] \end{aligned} \quad (209)$$

Using the relevant Taylor series expansions (see Eqs. (198)), we obtain

$$\partial_n \text{tr} K|_{n=0} = -\frac{1}{2} \text{tr}[\sigma_2] \quad (210)$$

which gives, correct to  $O(n^2)$ ,

$$R^A_{nA} = -\frac{1}{2} \text{tr}[\sigma_2] = R^A_{tA} \quad (211)$$

and

$$R^{AB}_{CD} = {}^{(2)}R^{AB}_{CD} + O(n^2)$$

which, in turn, implies

$$R^{AB}_{AB} = R_{||} + O(n^2) \quad (212)$$

Finally, we use the following general expression for the Einstein tensor (see, for example, [26], section 14.2, pp. 344):

$$\begin{aligned} G^t_t &= -\left(R^A_{nA} + \frac{1}{2}R^{AB}_{AB}\right) \\ G^n_n &= -\left(R^A_{tA} + \frac{1}{2}R^{AB}_{AB}\right) \end{aligned} \quad (213)$$

Plugging in the above expressions for the Riemann tensor, and finally taking the limit  $n \rightarrow 0$ , we immediately obtain part of Eq. (204). (Note that, with the up-down components,  $G_{\hat{\xi}}^{\hat{\xi}} = G_t^t$ .) One can go further and analyze, in the same way, the remaining components of the Einstein tensor; the final result is [110]

$$E_{\hat{b}}^{\hat{a}}|_H = \left[ \begin{array}{cc|c} E_{\perp} & 0 & 0 \\ 0 & E_{\perp} & 0 \\ \hline 0 & 0 & E_{\parallel}^{\hat{A}} \end{array} \right] \quad (214)$$

where  $E_{\hat{b}}^{\hat{a}}|_H = (16\pi)^{-1} G_{\hat{b}}^{\hat{a}}|_H$ .

One can now proceed to do a similar analysis and obtain the near-horizon symmetries for a generic Lanczos-Lovelock Lagrangian. The full result, turns out to be the same as Eq. (204). We skip the proof which involves a bit of combinatorics, but otherwise straightforward to establish. Here, we shall only give a proof of the identity  $E_{\hat{\xi}}^{\hat{\xi}} = E_{\hat{n}}^{\hat{n}}|_H$ , which would be directly relevant when we discuss thermodynamic structure of the field equations.

Using the first equality in Eq. (30), it is easy to deduce the form of  $E_k^k$  (no summation over  $k$ ). The alternating determinant simplifies upon using  $\delta_k^k = 1$ , and the fact that none of the other indices can be  $k$  due to total anti-symmetry. Therefore, we are left with

$$E_k^k = -\frac{1}{2} L_m \{\bar{k}\} \quad (215)$$

where  $L_m \{\bar{k}\}$  denotes terms which do not contain the index  $k$  at all. We now specialize to the case when the index  $k$  corresponds to  $n$ . The RHS can be further split depending on the number of occurrences of the index  $t$ , as

$$E_n^n = -\frac{1}{2} [L_m \{\bar{n}, t\} + L_m \{\bar{n}, 2t\} + L_m \{\bar{n}, \bar{t}\}] \quad (216)$$

The first set on the RHS contains terms like  $R_{AB}^{tD}$  which are identically zero (see first of Eqs. (207)), while the last set is the same as appears in  $E_t^t$ . So we only have to prove that  $L_m \{\bar{n}, 2t\} = L_m \{\bar{t}, 2n\}$ . The set  $L_m \{\bar{n}, 2t\}$  will have two  $t$ 's appearing either on *different* factors of  $R_{cd}^{ab}$ , which would again vanish identically for the same reason as the first set, or it can have the two  $t$ 's appearing on the *same* factor, which would contribute

$$\begin{aligned} L_m \{\bar{n}, 2t\} &= 4m \times \frac{1}{16\pi} \frac{1}{2^m} \delta_{tD_1 \dots C_m D_m}^{tB_1 \dots A_m B_m} R_{tB_1}^{tD_1} \dots R_{A_m B_m}^{C_m D_m} \\ &= 4m \times \frac{1}{16\pi} \frac{1}{2^m} \delta_{D_1 \dots C_m D_m}^{B_1 \dots A_m B_m} R_{tB_1}^{tD_1} \dots R_{A_m B_m}^{C_m D_m} \end{aligned} \quad (217)$$

Using Eqs. (209), we see that  $L_m \{\bar{n}, 2t\} = L_m \{\bar{t}, 2n\} + O(n^2)$ . Therefore,  $E_n^n = E_t^t + O(n^2)$ , and the equality holds on the horizon. The form for  $E_{\parallel \hat{B}}^{\hat{A}}$  maybe deduced in the same way, but since it is not of much use in study of horizon thermodynamics, we do not give that analysis here.

#### 4.1.2 Thermodynamic structure of Lanczos-Lovelock field equations

To demonstrate the equivalence of field equations with the identity  $TdS = dE + PdV$ , we shall use the results derived above in Sec. 4.1.1. We begin with the  $t - t$  component of the field equations, and, following an analysis similar to the one leading to Eq. (30) of [112], we arrive at,

$$E_{\hat{\xi}}^{\hat{\xi}} = E_t^t = \frac{1}{16\pi} \frac{m}{2^m} [\sigma_2]_{CB} \sigma^{CA} \mathcal{E}_A^B - \frac{1}{2} L_m^{(D-2)} + O(n) \quad (218)$$

where

$$\mathcal{E}_A^B = \delta_{AC_1 \dots D_{m-1}}^{BA_1 \dots B_{m-1}} {}^{(D-2)} R_{A_1 B_1}^{C_1 D_1} \dots {}^{(D-2)} R_{A_{m-1} B_{m-1}}^{C_{m-1} D_{m-1}} \quad (219)$$

In Sec. 4.1.1, we proved the equality  $E_{\hat{\xi}}^{\hat{\xi}} = E_{\hat{n}}^{\hat{n}}$  on the horizon. Therefore, we have

$$E_{\hat{\xi}}^{\hat{\xi}}|_H = E_{\hat{n}}^{\hat{n}}|_H = \frac{1}{16\pi} \frac{m}{2^m} [\sigma_2]_{CB} \sigma^{CA} \mathcal{E}_A^B - \frac{1}{2} L_m^{(D-2)}$$

Using again the Taylor series from Sec. 4.1.1, and the affine parameter  $\lambda$  introduced in that section, we also have

$$\delta_\lambda \sigma_{AB} = \frac{\delta \lambda}{\kappa} [\sigma_2]_{AB} + O[(\lambda - \lambda_H)^{1/2} \delta \lambda] \quad (220)$$

Using this, we arrive at the expression:

$$2E_{\hat{\xi}}^{\hat{\xi}} \sqrt{\sigma} \delta \lambda \stackrel{H}{=} T \underbrace{\left( \frac{1}{8} \frac{m}{2^{m-1}} \right) \mathcal{E}^{BC} \delta_\lambda \sigma_{BC} \sqrt{\sigma}}_{\delta_\lambda S} - \underbrace{\frac{L_m^{(D-2)} \sqrt{\sigma}}{2}}_{\delta_\lambda E} \delta \lambda \quad (221)$$

where we have introduced the Hawking temperature,  $T = \kappa/(2\pi)$ . As the final step, we will now show that the factor multiplying  $T$ , as well as the second factor on the right hand side, can both be expressed as variations of quantities defined in terms of intrinsic horizon geometry, as is anticipated by the notations “ $\delta_\lambda S$ ” and “ $\delta_\lambda E$ ” we have used for them. *This is a remarkable fact, and is most probably true only for the Lanczos-Lovelock models.* Moreover, the quantities  $S$  and  $E$  will turn out to be the entropy and energy associated with the horizon in Lanczos-Lovelock theory.

To begin with, consider the quantity

$$S = 4\pi m \int d\Sigma L_{m-1}^{(D-2)} \quad (222)$$

and note that, purely algebraically, the variation of the above expression must give equations of motion for the  $(m-1)^{\text{th}}$  order Lanczos-Lovelock term in  $(D-2)$  dimensions. (The variation would also produce surface terms, which would not contribute when evaluated at

$\lambda = \lambda_H$  because we will assume that the horizon is a compact surface with no boundary.) We have:

$$\delta_\lambda S = -4\pi m \int_H d\Sigma \mathcal{E}'^{BC} \delta_\lambda \sigma_{BC} \quad (223)$$

where we have evaluated the variation on  $\lambda = \lambda_H$ . Noting that the Lagrangian is

$$L_{m-1}^{(D-2)} = \frac{1}{16\pi} \frac{1}{2^{(m-1)}} \delta_{C_1 D_1 \dots D_{m-1}}^{A_1 B_1 \dots B_{m-1}} \dots (D-2) R_{A_{m-1} B_{m-1}}^{C_{m-1} D_{m-1}}$$

and using the first of Eqs. (30), we see that

$$\mathcal{E}'^B_C = -\frac{1}{2} \frac{1}{16\pi} \frac{1}{2^{(m-1)}} \mathcal{E}^B_C \quad (224)$$

Therefore, we obtain

$$\delta_\lambda S = \frac{1}{8} \frac{m}{2^{(m-1)}} \int_H d\Sigma \mathcal{E}^{BC} \delta_\lambda \sigma_{BC} \quad (225)$$

which is precisely the integral of the factor multiplying  $T$  in Eq. (221). As mentioned above,  $S$  defined in Eq. (222) is a function of  $\lambda$ , and its derivative with respect to  $\lambda$  is well defined and finite on the horizon. The expression for  $S$ , evaluated at  $\lambda = \lambda_H$ ,

$$S(\lambda = \lambda_H) = 4\pi m \int_H d\Sigma L_{m-1}^{(D-2)} \quad (226)$$

is what we shall interpret as the entropy of the horizon. As mentioned before,  $S$  has no physical significance away from the horizon. (As an aside, we would also like to point out the following fact; when  $D = 2m$ , that is, when the corresponding Lanczos-Lovelock action is the Euler characteristic of the full  $D$ -dimensional manifold, then  $D - 2 = 2(m - 1)$ , and we see from Eq. (226) that  $S$  is proportional to the Euler characteristic of the  $(D - 2)$ -dimensional horizon surface, determined by  $L_{m-1}^{(D-2)}$ ; therefore,  $S$  is just a constant number. For example, the Gauss-Bonnet term ( $m = 2$ ) in  $D = 4$  will contribute a constant additive term to the standard Bekenstein-Hawking entropy.)

Multiplying Eq. (221) by  $d^{(D-2)}y$ , integrating over the horizon surface, and taking the  $n \rightarrow 0$  limit, we now see that it can be written as

$$\begin{aligned} T \delta_\lambda S - \int_H d\Sigma L_m^{(D-2)} \delta\lambda &= \int_H d\Sigma T_{\hat{\xi}}^{\hat{\xi}} \delta\lambda \\ &= \int_H d\Sigma T_{\hat{n}}^{\hat{n}} \delta\lambda \\ &= \int_H d\Sigma P_{\perp} \delta\lambda \end{aligned} \quad (227)$$

where we have used the field equations  $E_{\hat{\xi}}^{\hat{\xi}} = (1/2)T_{\hat{\xi}}^{\hat{\xi}}$  in the first equality, and the relation  $E_{\hat{\xi}}^{\hat{\xi}}|_H = E_{\hat{n}}^{\hat{n}}|_H$  in the second equality. We have also used the interpretation of  $T_{\hat{n}}^{\hat{n}}$  as normal

pressure,  $P_\perp$ , on the horizon. The above equation now has the desired form of the first law of thermodynamics, provided: (i) We identify the quantity  $S$ , defined by Eq. (226) as the entropy of horizons in Lanczos-Lovelock gravity; indeed, *exactly* the same expression for entropy has been obtained in the literature using independent methods, see e.g, Ref.[75]. (That the same expression is obtained using the Noether charge approach of Wald, was demonstrated in Ref. [113].) (ii) We also identify the second term on left-hand-side as  $\delta_\lambda E$ ; this leads to the definition of  $E$  to be

$$E = \int^\lambda \delta\lambda \int_H d\Sigma L_m^{(D-2)} \quad (228)$$

where the  $\lambda \rightarrow \lambda_H$  limit must be taken *after* the integral is done (therefore, we would require the detailed form of  $L_m^{(D-2)}$  as a function of  $\lambda$  to calculate this explicitly). For  $D = 2(m+1)$ , the integral over  $H$  above is related to the Euler characteristic of the horizon, in which case  $E \propto \lambda_H$  (upon setting the arbitrary integration constant to zero). For  $m = 2, D = 4$ , this reduces, of course, to the expression obtained earlier in the case of Einstein gravity.

Let us briefly comment on the general form of  $E$  for *spherically symmetric* spacetimes for general Lanczos-Lovelock Lagrangians, with horizon at  $r = r_H$ , and  $\lambda = r$ . In this case,  $L_m^{(D-2)} \sim (1/\lambda^2)^m$  and also  $\sqrt{\sigma} \sim \lambda^{D-2}$ . The integrand in the energy expression therefore scales as  $\lambda^{(D-2)-2m}$ . Integrating the RHS of Eq. (228), and taking the  $\lambda \rightarrow \lambda_H$  limit after doing the integration, we see that  $E \sim \lambda_H^{(D-2)-2m+1}$ . As mentioned above, for  $D = 2(m+1)$ ,  $E \sim \lambda_H$ . In fact, in the case of spherically symmetric spacetimes in Lanczos-Lovelock theory, the above expression can be formally shown to be *exactly equivalent* to the one already known in the literature (see [112], and also Ref. [16] therein). As far as we are aware, no general expression for energy in Lanczos-Lovelock theory exists in the literature, and the one we have arrived at could be thought of as one definition which appears to be reasonable from physical point of view. In any case, the expression for energy clearly deserves further investigation.

Putting everything together, we see that, for *generic static spacetimes* in Lanczos-Lovelock gravity, the field equations can be written as a thermodynamic identity:

$$T\delta_\lambda S - \delta_\lambda E = \overline{F}\delta\lambda \quad (229)$$

where we have found it appropriate to interpret

$$\overline{F} = \int_H P_\perp \sqrt{\sigma} d^2y \quad (230)$$

as the the average normal force over the horizon “surface” (somewhat in line with the spirit of the so called *membrane paradigm*, which we discuss in further detail in Sec. 3.3). The term  $\overline{F} \delta\lambda$  has the interpretation of virtual work done in displacing the horizon by an affine distance  $\delta\lambda$ , in perfect analogy with the work term in the standard first law of thermodynamics.

We have therefore established the equivalence of Lanczos-Lovelock field equations with the thermodynamic identity, generalizing the result already known in Einstein gravity. In fact, as will be a recurring point in this review, we have achieved more than just a generalization of result from Einstein to all of Lanczos-Lovelock – the analysis here provides a more rigorous and robust geometric basis for the thermodynamic interpretation of field equations in Einstein



theory itself. To put the result in appropriate context, note that we have shown that the following relations

$$\underbrace{E(\hat{n}, \hat{n}) = -E(\mathbf{u}, \mathbf{u})}_{\text{near-horizon symmetry}} \quad \& \quad \underbrace{E(\mathbf{u}, \mathbf{u}) = (1/2) T(\mathbf{u}, \mathbf{u})}_{\text{field equations}}$$

have a thermodynamic structure. Democracy of all observers then implies that  $E = (1/2)T$  also has a thermodynamic structure:

$$E(\mathbf{u}, \mathbf{u}) = \frac{1}{2} T(\mathbf{u}, \mathbf{u}) \xrightarrow{\text{all } \mathbf{u}} E = \frac{1}{2} T$$

Further, in the limit  $N \rightarrow 0$ , we have  $N^2 \mathbf{k} \rightarrow \boldsymbol{\xi}$ . Using this and  $\mathbf{u} = \boldsymbol{\xi}/N$ , we see that  $T(\mathbf{u}, \mathbf{u}) \rightarrow T(\boldsymbol{\xi}, \mathbf{k})$ , which is the flow of energy for a virtual displacement of the horizon along  $\mathbf{k}$ . Another way to interpret this result is as follows: the object  $g_{ab}^\perp = 2\xi_{(a}k_{b)}$  is the transverse part of the induced metric, and hence  $T_{ab}\xi^a l^b = (1/2) \text{tr}_2[T_{ab}]$ , where  $\text{tr}_2$  is the trace with respect to the metric  $2\xi_{(a}k_{b)}$ . For spherically symmetric spacetimes, since  $T_0^0 = T_r^r = P$ , we have  $(1/2) \text{tr}_2[T_{ab}] = P$ , which is the result derived in Ref. [30, 114].

#### 4.1.3 Field equations from the Clausius relation

The above result suggests a strong relation between gravitational dynamics and horizon thermodynamics. This idea can be further strengthened by another result, viz. that gravitational field equations can be interpreted as the entropy balance condition on local Rindler horizons. In the most straight forward manner, this result can be obtained very simply from the following argument [115].

Consider an infinitesimal displacement of a local patch of the stretched (local Rindler) horizon  $\mathcal{H}$  along the direction of its normal  $r_a$ , by an infinitesimal proper distance  $\epsilon$ , which will change the proper volume by  $dV_{prop} = \epsilon \sqrt{\sigma} d^{D-2}x$  where  $\sigma_{ab}$  is the metric in the transverse space. The flux of matter energy through the surface will be  $T_b^a \xi^b r_a$  (where  $\xi^a$  be the approximate Killing vector corresponding to translation in the local Rindler time) and the corresponding entropy flux can be obtained by multiplying the energy flux by  $\beta_{loc} = N\beta$ . Hence the ‘loss’ of matter entropy perceived by the outside observer, because the virtual displacement of the horizon has engulfed some matter, is

$$\delta S_m = \beta_{loc} \delta E = \beta_{loc} T^{aj} \xi_a r_j dV_{prop}. \quad (231)$$

Interpreting  $\beta_{loc} J^a$ , where  $J^a$  is the Noether current corresponding to the local Killing vector  $\xi^a$  given by  $J^a = 2\mathcal{G}_b^a \xi^b + L\xi^a$ , as the gravitational entropy current, the change in the gravitational entropy is given by

$$\delta S_{\text{grav}} \equiv \beta_{loc} r_a J^a dV_{prop} \quad (232)$$

(Note the occurrence of the local, redshifted, temperature through  $\beta_{loc} = N\beta$  in these expressions.) As the stretched horizon approaches the true horizon,  $Nr^a \rightarrow \xi^a$  and  $\beta\xi^a \xi_a L \rightarrow 0$ . Hence we get, in this limit:  $\delta S_{\text{grav}} \equiv \beta \xi_a J^a dV_{prop} = 2\beta \mathcal{G}^{aj} \xi_a \xi_j dV_{prop}$ . Comparing  $\delta S_{\text{grav}}$  and  $\delta S_m$  we see that the field equations  $2\mathcal{G}_b^a = T_b^a$  can be interpreted as the entropy balance

condition  $\delta S_{\text{grav}} = \delta S_{\text{matt}}$ . This provides a direct thermodynamic interpretation of the field equations as local entropy balance in local Rindler frame.

Though we work with entropy density, the factor  $\beta = 2\pi/\kappa$  cancels out in this analysis — as it should, since the local Rindler observer with a specific  $\kappa$  was introduced only for interpretational convenience — and the relation  $T\delta S_m = T\delta S_{\text{grav}}$  would have the same interpretation. The expression in the right hand side is the change in the horizon (‘heat’) energy  $H_{\text{sur}} = TS$  due to injection of matter energy. (For a more detailed discussion of this expression, see Ref. [116]). The context we consider corresponds to treating the local Rindler horizon as a physical system (like a hot metal plate) at a given temperature and possessing certain intrinsic degrees of freedom. Then integration  $\delta S = \delta E/T$  at constant  $T$  to relate change in horizon energy to injected matter energy. Any energy injected onto a null surface hovers just outside the horizon for a very long time as far as the local Rindler observer is concerned and thermalizes at the temperature of the horizon if it is assumed to have been held fixed. This is a local version of the well known phenomenon that, the energy dropped into a Schwarzschild black hole horizon appears to hover just outside  $R = 2M$  as far as an outside observer is concerned. In the case of a local Rindler frame, similar effects will occur as long as the acceleration is sufficiently high; that is, if  $\kappa/\kappa^2 \ll 1$ .

While the above results hold quite rigorously, they only show that the field equations lead to an entropy balance condition *while what we are interested in is the converse*. We want to prove that the entropy balance condition implies the field equations. There has been attempts in the literature to ‘reverse-engineer’ the above argument [117] but these ideas face a major stumbling block. As pointed out by Padmanabhan [118] the reverse engineering requires interpreting  $J^a$  as an entropy current which in turn is justified only if we assume the field equations (e.g, Wald needed the field equations to show that Noether charge is physical entropy.)

Clearly the idea works only if we can have an expression for entropy *which can be independently justified*. In Einstein’s theory, entropy is proportional to area of the horizon which allowed Jacobson [119] to use the entropy balance condition to obtain the field equation in general relativity from Clausius relation. The result relied on three crucial inputs: (a) The entropy of the horizon is proportional to the area. (b) Raichaudhury equation relates the change in the area of cross-section of null congruence to the matter stress tensor. (c) Certain technical assumptions regarding the shear and expansion of the congruence. Given these, it is straight forward to obtain Einstein’s equations from the Clausius relation.

Various attempts have been made towards generalization of the original result of Jacobson to higher curvature gravity theories. Since entropy is no longer proportional to the area in a general Lanczos-Lovelock model, the Raichaudhury equation is pretty much useless as a tool and there is no way of directly generalizing the original idea of Jacobson. As a consequence, all the “generalizations” are either quite different from the original idea at a conceptual level (in that the Clausius relation is not employed in the way it was originally done), or else marred by technical errors and/or ad hoc assumptions tailor-made to get the desired result. The most immediate generalization given by Jacobson and collaborators themselves [120], for the  $f(R)$  theories, is distinct in concept in that additional dissipative terms had to be added to the heat flux to get the “correct” field equations. (Obviously, *any* field equation can be written in the form  $TdS - dE = (\text{something})$  with the right hand side interpreted as giving some ‘dissipation’, so this result is conceptually vacuous.) More recent attempts towards generalization are, on the other hand, based on a judicious use of the expression

for the Noether potential  $J_{ab}$ ; essentially, since diffeomorphism invariance of the action is a non-dynamical symmetry, the corresponding current  $J^a$  is conserved off-shell, and using its expression for a Killing field, one can hope to derive the equations of motion from a relation like  $TdS = \delta E$ . In these works, while the definition of matter flux is similar to Jacobson's, the Raychaudhuri equation is never invoked, thereby avoiding any need for some assumptions such as vanishing expansion etc. present in original Jacobson work.

The situation, however, is far from clear since these works do not all agree with each other. For e.g., Parikh and Sarkar have pointed out issues with the Brustein and Hadad [117] paper whereas [118] has highlighted certain conceptual issues regarding *interpretation* of the results. Specifically, [118] has stressed some of the subtleties (alluded to above) in interpreting these results as a *derivation* of field equations from thermodynamics; these results, as Ref. [118] argues, are better viewed as *interpretation* of field equations as a thermodynamic relation, rather than a derivation of the former from the latter. More importantly, Brustein and Hadad derive their result using different definitions for various quantities as compared to the other two papers; this calls for a more detailed and critical look at the analysis therein.

The most recent work in this direction is that in ref. [121] which we shall briefly describe. First, one starts with the following (very general) ansatz for horizon entropy

$$S = \int_{\Sigma} s^{ab} \epsilon_{ab} dA, \quad (233)$$

(where the entropy density  $s^{ab}$  is an antisymmetric tensor,  $\epsilon^{ab}$  is the binormal to the slice and  $dA$  is the area element on the slice), and then calculate the change of this entropy between two horizon slices  $\Sigma$  and  $\Sigma_0$  using Stokes' theorem

$$\delta S = \oint_{\Sigma \cup \Sigma_0} s^{ab} \epsilon_{ab} dA = -2 \int_H \nabla_b s^{ab} k_a dV dA \quad (234)$$

where  $V$  is a parameter along  $\mathbf{k}$  (which is tangent to horizon generators). Various assumptions about the slices and enclosed horizon region  $H$  can be found in [121]. Second, the Clausius relation,  $dS = \delta Q/T$ , is imposed with the heat flux given by

$$\delta Q = \int_H (-T_{ab}) \xi^b dH_a \quad (235)$$

Matching the order on two sides of the above relation appropriately using behaviour of the local Killing vector (LKV)  $\xi^j$  gives

$$-(\hbar/\pi) \nabla_b s^{ab} k_a = T^{ab} \xi_b k_a + O(x^2). \quad (236)$$

where  $O(x^2)$  refers to higher order terms in expansion of the LKV. Note that, at this stage, the relation derived above is nothing but the local entropy balance condition which was obtained earlier in [115], except that the form of  $s_{ab}$  has not yet been fixed. This third step is attempted as follows. One motivates a form for  $s_{ab}$  in terms of the LKV and its covariant derivatives, and imposes the above equation to determine conditions on (and inter-relationships between) the various (tensorial) coefficients in this expression. This leads to the final form for  $s_{ab}$  in terms of a 4 rank tensor with symmetries of Riemann, with certain integrability conditions imposed, which eventually identify this tensor with  $P_{abcd}$ . This basically summarizes the

key steps in [121] which attempts to generalize the Clausius relation maintaining as close a connection as possible with the original result of Jacobson. There are, however, key (and inevitable) differences from earlier work along these lines, and several other subtleties which are discussed in detail in [121].

Finally, we end this section by clarifying an important point concerning distinction between Sec. 4.1.3, where we discussed the equivalence of field equations with the first law of thermodynamics, and the discussion in the present section where one attempts to derive the field equations from the Clausius relation; that is, we wish to clarify the similarities and differences between  $dE + PdV$  and  $\delta Q$ , both of which are eventually equated to  $TdS$ . This was done in [122], where it was pointed out that:

1. There is a precise sense in which the “heat flux” term of Jacobson, and the “ $PdV$ ” term of Padmanabhan, are different; this difference is connected with the difference in null vectors which govern the horizon deformation in the two cases.
2. The change in area of a horizon cross-section is determined by the expansion (and not its first derivative) of the ingoing null congruence  $\mathbf{k}$  normalized to have unit Killing energy. An explicit expression for the expansion  $\theta$  of this congruence can be given in terms of combination of curvature tensor components as

$$\kappa \frac{d}{d\lambda} \ln \sqrt{\sigma} = R_{ab} \xi^a k^b - R_{abcd} u^a n^b u^c n^d \quad (237)$$

(where  $\lambda$  is the affine parameter along  $\mathbf{k}$ , and  $\mathbf{u}$  and  $\mathbf{n}$  are unit timelike and spacelike vectors spanning the horizon, see [122]), which clearly shows that *the area change involves not just Ricci, but also the Riemann tensor* – a point which is of relevance in the context of deriving field equations from thermodynamics.

3. From explicit  $(D - 2) + 1 + 1$  decomposition of the Einstein tensor, one can show that there is also a term corresponding to quasi-local energy of the horizon, which must be separately accounted for when considering energy flow across the horizon.
4. The Raychaudhuri equation with the prescribed null congruence has the  $O(\lambda)$  term which does not vindicate or necessitate setting the expansion to zero. The most curious aspect of the analysis, however, is the fact this term nevertheless yields the null-null component of Einstein equations, although the algebraic reason is completely different! In fact, in their recent generalization [121], Jacobson et. al. have indeed found that no conditions are imposed on their null congruence by the Clausius relation although this was invoked in the original paper of [119] and [120].

All these facts, along with the previous discussion in this section, shows that while *interpreting* the field equations as thermodynamic relations is mathematically rigorous, the *derivation* of field equations from thermodynamic considerations has unresolved conceptual and algebraic issues when it is attempted in the above manner. We shall see later (see Sec. 5) that this is best done from a completely different approach, viz. from a thermodynamic variational principle.

## 4.2 Equipartition law for microscopic degrees of freedom

Given the fact that spacetime appears to be hot, just like a body of gas, we can apply the Boltzmann paradigm (“If you can heat it, it has microstructure”) and study the nature of the microscopic degrees of freedom of the spacetime — exactly the way people studied gas dynamics *before* the atomic structure of matter was understood. One key relation in such an approach is the equipartition law  $\Delta E = (1/2)k_B T \Delta N$  relating the number density  $\Delta N$  of microscopic degrees of freedom we need to store an energy  $\Delta E$  at temperature  $T$ . (This number is closely related to the Avogadro number of a gas, which was known even before one figured out what it was counting!).

If gravity is the thermodynamic limit of the underlying statistical mechanics, describing the ‘atoms of spacetime’, we should be able to relate  $E$  and  $T$  of a given spacetime and determine the number density of microscopic degrees of freedom of the spacetime when everything is static. Remarkably enough, we can do this directly from the gravitational field equations [123, 124, 125]. The establishment of such a relation would obviously be important from the perspective of the emergent gravity paradigm as well.

Since the basic idea here is quite simple (it is explained, for e.g., [123, 124, 125]), it is convenient to first describe the result for Einstein theory (which is essentially some well known relations interpreted differently), and then describe the steps needed to derive it more formally for the entire class of Lanczos-Lovelock models using the Noether current and potential. We shall work with a static spacetime with metric components  $g_{00} = -N^2$ ,  $g_{0\alpha} = 0$ ,  $g_{\alpha\beta} = h_{\alpha\beta}$ . An observer at rest in this spacetime with four-velocity  $u^a = \delta_0^a/N$  will have an acceleration  $a^j = (0, a^\mu)$  where  $a_\mu = (\partial_\mu N/N)$ . We consider a spatial 3-volume  $\mathcal{V}$  in this spacetime, having a boundary  $\partial\mathcal{V}$ .

### 4.2.1 Equipartition law in Einstein theory

To fix the ideas, we will begin with the simplest context in which one can determine the density of microscopic degrees of freedom, viz., a static spacetime in Einstein’s general relativity. In a static spacetime, it is easy to show that

$$R_{ab}u^a u^b = \nabla_i a^i = N^{-1} D_\mu (N a^\mu) \quad (238)$$

where  $D_\mu$  is the covariant derivative operator corresponding to the 3-space metric  $h_{\alpha\beta}$ . Using the Einstein field equations in Eq. (238), we can relate the divergence of the acceleration to the source  $\bar{T}_{ab} \equiv (T_{ab} - (1/2)g_{ab}T)$  by:

$$D_\mu (N a^\mu) \equiv 8\pi N \bar{T}_{ab} u^a u^b \equiv 4\pi \rho_{\text{Komar}} \quad (239)$$

where  $\rho_{\text{Komar}}$  is the (so called) Komar mass-energy density. Integrating both sides of Eq. (239) over  $\mathcal{V}$  and using Gauss theorem gives:

$$E \equiv \int_{\mathcal{V}} d^3x \sqrt{h} \rho_{\text{Komar}} = \frac{1}{2} \int_{\partial\mathcal{V}} \frac{\sqrt{\sigma} d^2x}{L_P^2} \left\{ \frac{N a^\mu n_\mu}{2\pi} \right\} \quad (240)$$

where  $\sigma$  is the determinant of the induced metric on  $\partial\mathcal{V}$  and  $n_\mu$  is the spatial normal to  $\partial\mathcal{V}$ . (We have temporarily restored  $G = L_P^2 \neq 1$  keeping  $\hbar = c = k_B = 1$ .)

This result has a remarkable interpretation directly related to the equipartition law of conventional thermodynamics[123, 124, 125]. To see this, choose  $\partial\mathcal{V}$  to be a  $N = \text{constant}$  surface so that the normal  $n_\mu$  is in the direction of the acceleration and  $a^\mu n_\mu = |\mathbf{a}|$  is the magnitude of the acceleration. The local, redshifted, Davies-Unruh temperature[39] is then given by:  $T = NT_{\text{loc}} = (Na^\mu n_\mu/2\pi) = (N|\mathbf{a}|/2\pi)$ . Therefore Eq. (240) is exactly in the form of the equipartition law (with  $k_B = 1$ ) where

$$E = \frac{1}{2} \int_{\partial\mathcal{V}} dnT \quad ; \quad \Delta n \equiv \frac{\sqrt{\sigma} d^2x}{L_P^2} \quad (241)$$

That is, demanding the validity of Einstein equations, and interpreting Eq. (241) as a law of equipartition, we can *determine* the number of microscopic degrees of freedom in an element of area  $\Delta A = \sqrt{\sigma} d^2x$  on  $\partial\mathcal{V}$  to be  $\sqrt{\sigma} d^2x / L_P^2$ . If these microscopic degrees of freedom are in equilibrium at the temperature  $T$ , then the total equipartition energy contributed by all these degrees of freedom on  $\partial\mathcal{V}$  is equal to the total energy contained in the bulk volume enclosed by the surface — which could be thought of as another occurrence of the holographic principle.

#### 4.2.2 Equipartition law in Lanczos-Lovelock models

Remarkably enough, the equipartition law exists for *any* diffeomorphism invariant theory of gravity and allows us to identify the corresponding surface density  $\Delta n / \sqrt{\sigma} d^2x$  of microscopic states in a given theory. However, we restrict our attention to Lanczos-Lovelock models since the general case requires an additional assumption involving bifurcate Killing horizon which makes it's status a bit unclear.

To demonstrate how the result generalizes, we follow precisely the same steps [124, 125] as in the Einstein case, but now giving a generalized version of the equations encountered before. For Lanczos-Lovelock actions, the object which replaces Ricci tensor of Einstein theory is defined by  $E_{ab} + (1/2)Lg_{ab}$ , which, from Eq. (10), gives

$$\mathcal{R}_{ab} = P_a{}^{cde} R_{bcde} \quad (242)$$

(Note that  $\mathcal{R}_{ab} \Rightarrow (16\pi)^{-1} R_{ab}$  in the Einstein case.) We shall also require the expression for Noether current  $J^a$  associated with the timelike Killing vector  $\xi^a$ , which is quite simple and is given by  $J^a = 2\mathcal{R}_b^a \xi^b$ . One is now set to rerun the steps outlined for Einstein theory above, as follows.

Using the definition of  $J^{ab}$  (or expression Eq. (41) directly for a Killing vector), we get:

$$2\mathcal{R}_{ab} u^a u^b = \nabla_a (J^{ba} u_b N^{-1}) \quad (243)$$

Further, in a static spacetime,  $\nabla_i Q^i = N^{-1} D_\alpha (N Q^\alpha)$  for any static vector (with  $\partial_0 Q^i = 0$ ) and hence we can write this as:

$$2N\mathcal{R}_{ab} u^a u^b = D_\alpha (J^{b\alpha} u_b) \quad (244)$$

which generalizes of Eq. (238) to an arbitrary theory of gravity. The source for gravity in a general theory (analogous to Komar energy density) is *defined* through

$$\rho \equiv 4N\mathcal{R}_{ab} u^a u^b \quad (245)$$

For Lanczos-Lovelock models,  $\rho$  can be written as a function the matter stress-energy tensor  $T_{ab}$  (generalizing the  $\bar{T}_{ab}$  used to define Komar energy in Einstein case above) using the field equations. This is possible for Lanczos-Lovelock models since trace of the equations of motion is proportional to the Lagrangian. However, the explicit form will not be needed here. Integrating both sides over a region  $\mathcal{V}$  bounded by  $\partial\mathcal{V}$ , we get

$$\begin{aligned} E &\equiv \int_{\mathcal{V}} \sqrt{h} d^{D-1}x \rho \\ &= \int_{\mathcal{V}} 2N \mathcal{R}_{ab} u^a u^b \sqrt{h} d^{D-1}x \\ &= \int_{\partial\mathcal{V}} d^{D-2}x \sqrt{\sigma} (n_i u_b J^{bi}) = \int_{\partial\mathcal{V}} d^{D-2}x \sqrt{\sigma} (N n_\alpha J^{\alpha 0}) \end{aligned} \quad (246)$$

which is the analogue of Eq. (240). In Einstein's theory,  $J^{\mu 0} = a^\mu / 8\pi$  which will reduce Eq. (246) to Eq. (240). In a general theory, the expression for  $\Delta n$  is not just proportional to the area and  $J^{\alpha 0}$  quantifies this difference.

To obtain the equipartition law in a more explicit form, we note that (from Eq. (41)), we get  $J^{ab} = 2P^{abcd} \nabla_c \xi_d$  and hence  $J^{\alpha 0} = 4|\mathbf{a}| P_{0\beta}^{\alpha 0} n^\beta$ . So

$$E = \int_{\partial\mathcal{V}} d^{D-2}x \sqrt{\sigma} (16\pi P_{0\beta}^{\alpha 0} n^\beta n_\alpha) \left( \frac{N|\mathbf{a}|}{2\pi} \right) \equiv \frac{1}{2} \int dnT \quad (247)$$

where  $T = N|\mathbf{a}|/2\pi$  is the Davies-Unruh temperature as before. The surface density of microscopic degrees of freedom, when the field equations are satisfied (step 2 above), can be now read-off as:

$$\frac{dn}{\sqrt{\sigma} d^{D-2}x} = 32\pi P_{cd}^{ab} \epsilon_{ab} \epsilon^{cd} \quad (248)$$

where  $\epsilon_{ab} \equiv (1/2)(u_a n_b - u_b n_a)$  is the binormal to  $\partial\mathcal{V}$ . The connection with Wald entropy (see section (3.1)) is evident. In fact,

$$S = (1/2)\beta E \quad (249)$$

provides the relationship between entropy, energy and temperature in all Lanczos-Lovelock models including general relativity as a special case. (In Ref. [123], the implications of this result for general relativity was discussed extensively.) It also follows that, in all these models, we have  $4\Delta S = \Delta n$ .

To sum up, the equipartition law allows one to identify the number density of microscopic degrees of freedom on a constant redshift surface. Using this one can define the entropy of the horizon in a general theory of gravity, which, as shown above, agrees with Wald entropy. These ideas will play a key role in our discussion in Sec. 5. We now discuss one specific context in which the equipartition idea presented here might have some fundamental implication. However, what follows is mathematically independent of the results derived above.

### 4.2.3 Holography and quantization of gravitational entropy

Most of the arguments in the emergent gravity paradigm are free of unnecessary speculations since the approach taken is “top-down” and relies on well-established principles. For example,



the key result of the previous section expressed in the form of Eq. (248), was derived from a classical theory with the only quantum mechanical input being the formula for Davies-Unruh temperature. The  $\hbar$  enters the expression only through the formula  $k_b T = (\hbar/c)(a/2\pi)$  which, along with  $G$  in *classical* gravitational field equations leads to  $G\hbar/c^3$ , one of the central constants in *quantum* gravity. So one does not have to make any unwarranted speculations from the domain of quantum gravity to obtain Eq. (248).

However, it is worth exploring where a result like Eq. (248) might emerge from a microscopic candidate theory of quantum gravity like for example, string theory or loop quantum gravity. To begin with, it is clear that *our result in Eq. (248) should not depend on the details of the correct theory of quantum gravity*. Any model of quantum gravity which has (i) the correct classical limit and (ii) is consistent with Davies-Unruh temperature, will lead to our Eq. (248). Since one would expect any quantum gravity model to satisfy these two conditions (i) and (ii) above our “top-down” approach will meet the “bottom-up” approach of such a theory in the overlap domain. So if string theory or loop-quantum-gravity or dynamical triangulations or any other candidate model for quantum gravity satisfies (i) and (ii), it will lead to the equipartition. (On the other hand, if a quantum gravity model does not satisfy (i) and (ii), such a model is probably wrong in any case.) Even though the thermodynamic limit contains far less information about the system than a microscopic description, one could still wonder about the possible route or mechanism by which a *wide class* of candidate models in quantum gravity (all satisfying (i) and (ii) above) can lead to a relation like Eq. (248) in the appropriate limit. In this section, we will indicate one route through which such results can emerge. (Although the motivation for our discussion comes from ideas in the emergent gravity paradigm, the results themselves are of independent interest since they generalize certain ideas about quantum mechanical spectrum of a black hole and its connection with the highly damped quasi-normal modes.)

Let us first consider Einstein’s theory in which the equipartition law assigns  $A/L_P^2$  degrees of freedom to a proper area element  $A$  so that  $A \approx nL_P^2$  for  $n \gg 1$ . (This corresponds to the semiclassical limit in which one can meaningfully talk about proper area elements etc. in the semiclassical metric.) This is, of course, nothing but area quantization in the asymptotic limit. Though LQG probably leads to such a result, in the limit we are interested in, it can be obtained from the Bohr-Sommerfeld quantization condition applied to horizons. In fact, Bekenstein conjectured [126] years back that, in a quantum theory, the black hole *area* would be represented by a quantum operator with a discrete spectrum of eigenvalues. (Bekenstein showed that the area of a classical black hole behaves like an adiabatic invariant, and hence, according to Ehrenfest’s theorem, the corresponding quantum operator must have a discrete spectrum.) Extending these ideas to local Rindler horizons treated as the limit of a stretched horizon, one can understand how a result like quantization of any spatial area element can arise in a microscopic theory.

The situation becomes more interesting when we go from the Einstein gravity to Lanczos-Lovelock models. In Einstein gravity, entropy of the horizon is proportional to its area and hence one could equivalently claim that it is the *entropy* which has an equidistant spectrum. But, in the Lanczos-Lovelock models, this proportionality between area and entropy breaks down. (Generalization of LQG to Lanczos-Lovelock models does not exist as yet.) The question then arises as to whether it is the quantum of area or quantum of entropy (if either) which arises in these models. This question was addressed in ref. [127] where it was shown that in the Lanczos-Lovelock models, it is indeed *the entropy that is quantized* with an



equidistant spectrum. This matches nicely with the fact that it is the quantity in the right hand side of Eq. (248) that takes integral values in the equipartition law. Thus one can alternatively interpret our result as entropy quantization.

We now indicate a sufficiently general ‘mechanism’ by which any microscopic model of quantum gravity can lead to such a quantization condition, in terms of two features. First one is the peculiar ‘holographic’ structure of the action functionals in Lanczos-Lovelock models [128, 29, 21]. As noted in Sec. ...., the action functionals in *all* these theories can be separated into a bulk and surface term and the surface term has the structure of an integral over  $d(pq)$ . It can also be shown [3, 128, 29] that the same ‘ $d(pq)$ ’ structure emerges for the on-shell action in the Lanczos-Lovelock models; that is,

$$A|_{\text{on-shell}} \propto \int_V d^D x \partial_i (\Pi^{ijk} g_{jk}) = \int_{\partial V} d\Sigma_i \Pi^{ijk} g_{jk} \quad (250)$$

where  $\Pi^{ijk}$  is the suitably defined canonical momenta corresponding to  $g_{jk}$ . The second ingredient is that, in all these theories, the above expression for the action leads to the Wald entropy of the horizon. (For example, in Einstein gravity, the surface term in action, evaluated on a horizon will give one quarter of area [128, 29].) In the semiclassical limit, Bohr-Sommerfeld quantization condition demands that the integral of  $d(pq)$  should be equal to  $2\pi n$ . Since the action in Eq. (250) has this ‘ $d(pq)$ ’ structure, the Bohr-Sommerfeld condition reduces to  $A|_{\text{on-shell}} = 2\pi n$ . Since  $A|_{\text{on-shell}}$  is also equal to Wald entropy, we obtain

$$S_{\text{Wald}} = A|_{\text{on-shell}} = 2\pi n \quad (251)$$

The Bohr-Sommerfeld quantization condition, of course, was the same used originally by Bekenstein and others to argue for the *area* quantization of the horizon but in the more general context of Lanczos-Lovelock models it appears as entropy quantization [127]. (It is also possible to argue that in the *semiclassical* limit, the on-shell value of the action will be related to the phase of the semiclassical wave function  $\Psi \propto \exp(iA|_{\text{on-shell}})$ . If the semiclassical wave function describing the quantum geometry, relevant for a local Rindler observer, is obtained by integrating out the inaccessible degrees of freedom beyond the horizon then one can argue that [127], this phase should be  $2\pi n$  in the asymptotic limit leading to  $A|_{\text{on-shell}} = 2\pi n$ . Given the conceptual ambiguities related to interpretation of ‘wave function of geometry’, it is probably better to invoke Bohr-Sommerfeld condition.) To summarize, the three features: (i) the *holographic* structure of the gravitational action functional (ii) the equality of on-shell gravitational action functional and Wald entropy and (iii) the Bohr-Sommerfeld quantization condition, combine together to provide a context in which Eq. (248) can arise in any microscopic theory.

While this gives a general result that in Lanczos-Lovelock theories it is the entropy of the horizon that is quantized, it would be useful to reinforce it by an explicit calculation in the standard context. Fortunately, this can be done for the  $m = 2$  Lanczos-Lovelock term, that is, for the GB model, using arguments first suggested by Hod [129] based on quasi-normal modes of black hole oscillations. We give here a qualitative sketch of the arguments, and refer to [127] for technical details.

We first briefly summarize Hod’s argument. He started from Bekenstein’s arguments regarding quantum area spectrum of a non extremal Kerr-Newman black hole, and showed that the spacing of area eigenvalues can be fixed by associating the classical limit of the quasi-normal mode frequencies,  $\omega_c$ , with the large  $n$  limit of the quantum area spectrum, in the

spirit of Bohr's correspondence principle ( $n$  being the quantum number). For a Schwarzschild black hole of mass  $M$  in  $D = 4$  dimensions, the absorption of a quantum of energy  $\omega_c$ , would lead to change in the black hole area eigenvalues as,  $\Delta\mathcal{A} \equiv \mathcal{A}_{n+1} - \mathcal{A}_n = (\partial\mathcal{A}/\partial M)\omega_c$  and for entropy  $\Delta S = (\partial S/\partial M)\omega_c$ . For a Schwarzschild black hole the level spacing of both area and entropy eigenvalues were indeed found to be equidistant, allowing one to associate the notion of a minimum unit, *a quantum*, of area *and* entropy.

To apply this idea to Lanczos-Lovelock models, we need an expression for the quasi-normal mode frequencies associated with black hole solutions in these theories. These are not known in general, but are known in the case of  $5D$  Einstein-Gauss-Bonnet (GB) black holes. The form of the highly damped quasi-normal modes of these black holes suggest that it is the entropy which has a equally spaced spectrum [127]. The result essentially depends only on the fact that  $\text{Re } \omega_c \propto T$  leading to  $(\partial S/\partial M)\omega_c \propto T(\partial S/\partial M)$  which is a constant. On the other hand, since in this case area is a non-linear function of entropy,  $\mathcal{A} = F(S)$ , the area spectrum is *not* equidistant. Qualitative comments concerning the precise value of the quantum of entropy (in particular, the condition under which it is  $2\pi$ ) can be found in [127].

The case we have made for the quantization of entropy can be analyzed in several different ways. If one applies our argument based on quasi-normal frequencies with the *imaginary* rather than real part of the frequencies (a strong case for doing so was suggested by Maggiore [130]), then the result for constant spacing of entropy eigenvalues follows simply from the fact that this imaginary part, being derivable from a scattering matrix approach [131, 132], is independent of the specific model of gravity and depends only on the existence of a horizon.

The physical implications of the statement: *gravitational entropy is quantized*, are unclear at present. As should have been evident from our discussion in earlier sections, there are various proposed explanations of horizon entropy in the literature, with widely differing interpretations. For Lanczos-Lovelock models, most of these reproduce Wald entropy or yield expressions closely resembling it (as in Sec. 3.2), and therefore the statement that Wald entropy is quantized with equidistant spacing will also generically have implications for the physical principles on which a particular approach to horizon entropy is based. Since at least some of these approaches remain close to well known physics (for instance, the approach described in Sec. 3.2), the statement of quantization of entropy might be of relevance at more fundamental level than has yet been appreciated.

### 4.3 Emergent aspects of the Lanczos-Lovelock action

We now turn to the thermodynamic relevance of results discussed in Sec. 2.7 about the peculiar structure of the action functional and related issues. We saw that the splitting of the total action into two terms led to a holographic relationship between the two. It turns out that, there exists *another* splitting of Lanczos-Lovelock actions which *also* displays similar holographic structure. This separation is a generalization of the split of the Einstein-Hilbert Lagrangian as  $R = 2(G_0^0 - R_0^0)$ . For static spacetime with a Killing horizon, the 0-th component is uniquely determined by the Killing time and hence unambiguous. In the case of Einstein's theory, the first term ( $G_0^0$ ) in the above split is related to energy  $E$  while the second one ( $R_0^0$ ) to entropy  $S$ , of a (static) horizon. More precisely, the first term is associated with the ADM Hamiltonian while the second one to projection of the Noether charge corresponding to diffeomorphism invariance of the theory, which gives the entropy of horizon in such theories. This gives the total action a thermodynamic interpretation as the

free energy (at least in spacetimes with horizon). Further, it can be shown that the new pair above also *obeys the same holographic relation* Eq. (77); the surface term in the pair arises from the  $R_0^0$  part of the split - a result which holds for static spacetimes [21]. Since the bulk and surface terms in this case are related to energy and entropy, the holographic relation between them may be thought of as analogous to inverting the expression for entropy  $S = S(E, \dots)$  to obtain the energy  $E = E(S, \dots)$  in terms of entropy in normal thermodynamic systems.

It can be shown that the very same holographic relationship between the bulk and the surface term exists for an *arbitrary* splitting of the Lagrangian into  $L_{\text{sur}}$  and  $L_{\text{bulk}}$ , provided the trace of the equations of motion is proportional to the Lagrangian:  $g^{ab}E_{ab} \propto L$ . This is certainly true for Lanczos-Lovelock Lagrangians, although there is no rigorous proof as to whether there are any other Lagrangians satisfying it. (The proofs for the above statements can be found in [21].) We will now discuss these features.

#### 4.3.1 Lanczos-Lovelock action as free energy of spacetime

We now present one more aspect of the gravitational action functional which adds strength to the thermodynamic interpretation. One can show that, in any static spacetime with a bifurcation horizon (so that the metric is periodic with a period  $\beta = 2\pi/\kappa$  in the Euclidean sector), the *action functional for gravity can be interpreted as the free energy of spacetime* for all Lanczos-Lovelock models. The Noether current for Lanczos-Lovelock theory

$$J^a \equiv (2E^{ab}\xi_b + L\xi^a). \quad (252)$$

will play a prominent role in our discussion, where we have assumed staticity (so that  $\delta v^a = 0$ ), and will work in the natural coordinate system in which the timelike Killing vector field  $\xi^a = (1, \mathbf{0})$  explicitly exhibits the static nature.

We start with the Noether current  $J^a$  and integrate it on-shell (i.e., with  $2E_{ab} = T_{ab}$ ), over a constant- $t$  hypersurface with the measure  $d\Sigma_a = \delta_a^0 N \sqrt{h} d^{D-1}x$  where  $g_{00}^E = N^2$  and  $h$  is the determinant of the spatial metric. (We will work in the Euclidean sector.) Multiplying by the period  $\beta$  of the imaginary time, we get

$$\begin{aligned} \beta \int J^a d\Sigma_a &= \beta \int T_b^a \xi^b d\Sigma_a + \beta \int L \xi^a d\Sigma_a \\ &= \int (\beta N) T_b^a u_a u^b \sqrt{h} d^{D-1}x + \int_0^\beta dt_E \int L \sqrt{g} d^{D-1}x \end{aligned} \quad (253)$$

where  $u^a = \xi^a/N = N^{-1}\delta_0^a$  is the 4-velocity of observers moving along the orbits of  $\xi^a$  and  $d\Sigma_a = u_a \sqrt{h} d^{D-1}x$ . The term involving the Lagrangian gives the Euclidean action for the theory, while in the term involving  $T_{ab}$ , we note that  $\beta N \equiv \beta_{\text{loc}}$  is the redshifted local temperature. If we define the (thermally averaged) energy  $E$  as

$$\int (\beta N) T_b^a u_a u^b \sqrt{h} d^{D-1}x = \int \beta_{\text{loc}} T_b^a u_a u^b \sqrt{h} d^{D-1}x \equiv \beta E \quad (254)$$

we obtain

$$A = \beta \int J^a d\Sigma_a - \beta E \quad (255)$$

Upon using Stokes theorem, we see that the first term (with the Noether charge) gives horizon entropy. Therefore, we find that

$$A = S - \beta E = -\beta F \quad (256)$$

where  $F$  is the free energy. (Usually, one defines the Euclidean action with an extra minus sign as  $A = -A_E$  so that the Euclidean action can be interpreted directly as the free energy.) One can also obtain from Eq. (253) the relation:

$$S = \beta \int J^a d\Sigma_a = \int \sqrt{-g} d^D x (\rho + L) \quad (257)$$

where  $\rho = T_{ab} u^a u^b$ . This gives the entropy in terms of matter energy density and the Euclidean action. Alternatively, if we assume that the Euclidean action can be interpreted as the free energy, then these relations provide an alternative argument for interpreting the Noether charge as the entropy. (Similar results have been obtained in Ref. [42, 43] by a more complicated procedure.)

There is another way of expressing the above result by introducing a vector  $l_k \equiv \nabla_k t = \partial_k t = (1, \mathbf{0})$  which is (un-normalized) normal to  $t = \text{constant}$  hypersurfaces. The unit normal is  $\hat{l}_k = N l_k = -u_k$  in the Lorentzian sector. Using  $J^a = \nabla_b J^{ab}$ , we can write

$$2 E_b^k \xi^b + L \xi^k = J^k = \nabla_a J^{ka} \quad (258)$$

and since  $l_k \nabla_a J^{ka} = \nabla_a (l_k J^{ka})$ , we get

$$L = \nabla_a (l_k J^{ka}) - 2 E_b^k \xi^b l_k = \frac{1}{\sqrt{-g}} \partial_\alpha (\sqrt{-g} l_k J^{k\alpha}) - 2 E_b^k \xi^b l_k \quad (259)$$

which is equivalent to the previous expression with  $J^a$ , but now expressed in a more formal way and using Noether potential (rather than current).

The free-energy structure of the Lanczos-Lovelock action can be very explicitly demonstrated for black hole solutions which are *spherically symmetric*, in which case, the the action can be written as a total derivative. The details can be found in section 4 of Ref.[60], and follows from an algebraic identity which holds for Lanczos-Lovelock Lagrangians of a given order  $m$  evaluated for spherically symmetric spacetimes with  $-g_{00} = f(r) = g^{rr}$  (with horizon at  $f(r=a)=0$ ):

$$r^{D-2} L_m = \frac{d}{dr} \left[ \frac{(D-2)!}{(D-2m-1)!} (1-f)^m r^{D-2m-1} \right] - \frac{d}{dr} \left[ \frac{(D-2)!}{(D-2m)!} m f' (1-f)^{m-1} r^{D-2m} \right] \quad (260)$$

The Euclidean action, given by  $\int_0^\beta d\tau \int d^3 x \sqrt{g_E} L_m[f(r)]$ , when integrated over a spacetime

region bounded by the horizon at  $r = a$  and an arbitrary radius  $r = b(> a)$ , becomes

$$\begin{aligned}
-A_{\text{e } m} &= \frac{\beta\Omega}{16\pi} \int_a^b dr \left\{ \left[ \frac{(D-2)!}{(D-2m-1)!} (1-f)^m r^{D-2m-1} \right]' - \left[ \frac{(D-2)!}{(D-2m)!} m f' (1-f)^{m-1} r^{D-2m} \right]' \right\} \\
&= \frac{\beta\Omega}{16\pi} \left[ \underbrace{\frac{(D-2)!}{(D-2m-1)!} a^{D-2m-1}}_{(16\pi/\Omega)E_m} - (4\pi T) \underbrace{\frac{m(D-2)!}{(D-2m)!} a^{D-2m}}_{4S_m/\Omega} \right] + Q[f(b), f'(b)] \\
&= \beta E_m - S_m + Q[f(b), f'(b)]
\end{aligned} \tag{261}$$

where  $\Omega$  is the area of a unit  $(D-2)$ -sphere, and we have used known expressions for horizon energy and entropy for spherically symmetric black holes in Lanczos-Lovelock theory. (The functional  $Q$  depends on behaviour of the metric near  $r = b$ , and is irrelevant for our argument.) The free energy structure is self-evident in the contribution to the Euclidean action on the horizon, thereby supporting the general arguments given earlier in this section.

## 5 Entropy density of spacetime and Lanczos-Lovelock models

As already described in the previous section, one of the most important basis for treating gravity as an emergent phenomenon is the possibility of deducing field equations of gravity from conventional laws of thermodynamics applied to spacetime horizons perceived by relevant class of observers. Since horizons play a central role in such discussions, it is natural to consider an approach in which null surfaces provide the fundamental variables from which gravitational dynamics of the background spacetime emerges. Indeed, since null surfaces act as one-way membranes and can block information for a specific family of observers (time-like curves), it is natural to attribute an entropy to such surfaces.

If we take this point of view seriously, then the deformations of spacetime  $(\bar{x}^i - x^i) \equiv q^i$  associated with a vector field  $q^i$  becomes analogous to deformations of a solid in the study of elasticity. By and large, such a spacetime deformation is not of much consequence except for the null surfaces. As we have described earlier, null surfaces can be thought of as acting as a local Rindler horizon to a suitable set of observers. The deformation of a local patch of a null surface will change the amount of information accessible to the local Rindler observer. Equivalently, such an observer will associate certain amount of entropy density with the deformation of a null patch with normal  $n^a$ . This suggests that extremizing the sum of gravitational and matter entropy associated with *all* null vector fields *simultaneously*, could lead to the equations obeyed by the background metric.

Conceptually, this idea is similar to the manner in which we determine the influence of gravity on other matter fields. If we fill the spacetime with freely falling observers and insist that normal laws of special relativity should hold for all inertial observers simultaneously, we can arrive at the generally covariant versions of equation obeyed by matter in an arbitrary metric. This, in turn, allows us to determine the influence of gravity on matter fields and fixes the *kinematics* of gravity. To determine the *dynamics*, we do the same but now by filling

the spacetime with local Rindler observers. Insisting that the local thermodynamics should lead to extremum of an entropy functional associated with every null vector in the spacetime, we should be able to obtain the equations which determine the background spacetime.

*There is no a priori assurance that such a program will succeed* and hence it is a surprise that one can actually achieve this. Let us associate with every null vector field  $n^a(x)$  in the spacetime a thermodynamic potential  $\mathfrak{S}(n^a)$  (say, entropy) which is quadratic in  $n^a$  and given by:

$$\mathfrak{S}[n^a] = \mathfrak{S}_{grav}[n^a] + \mathfrak{S}_{matt}[n^a] \equiv - (4P_{ab}^{cd} \nabla_c n^a \nabla_d n^b - T_{ab} n^a n^b) , \quad (262)$$

where  $P_{ab}^{cd}$  and  $T_{ab}$  are two tensors that play a role analogous to elastic constants in the theory of elastic deformations. If we extremize this expression with respect to  $n^a$ , one will get a differential equation for  $n^a$  involving its second derivatives. We, however, want to demand that the extremum holds for all  $n^a$ , thereby constraining the *background* geometry. Further, the insistence on strictly local description of null-surface thermodynamics translates into the demand that the Euler derivative of the functional  $\mathfrak{S}(n^a)$  should not contain any derivatives of  $n^a$ .

It is indeed possible to satisfy these conditions by the following choice: We take  $P_{ab}^{cd}$  to be a tensor having the symmetries of curvature tensor and divergence-free in all its indices and take  $T_{ab}$  to be a divergence-free symmetric tensor. (The conditions  $\nabla_a P_{cd}^{ab} = 0$ ,  $\nabla_a T_b^a = 0$  can be thought of as a generalization of the notion of “constancy” of elastic constants of spacetime [133].) Once we get the field equations we can read off  $T_{ab}$  as the matter energy-momentum tensor; our notation anticipates this result. We also know that the  $P^{abcd}$  with the assigned properties can be expressed as  $P_{ab}^{cd} = \partial L / \partial R_{cd}^{ab}$  where  $L$  is the Lanczos-Lovelock Lagrangian and  $R_{abcd}$  is the curvature tensor [3, 4]. (Again, the notation already anticipates the fact that it will turn out to be the same object as we have introduced in earlier parts of this review. However, since we are now hoping to arrive at the Lanczos-Lovelock actions in the context of emergent gravity paradigm, we hope to present a new perspective on the physical relevance of this tensor.) This choice in Eq. (262) will also ensure that the equations resulting from the entropy extremisation do not contain derivatives of the metric of higher order than second.

We now demand that  $\delta \mathfrak{S} / \delta n^a = 0$  for the variation of all null vectors  $n^a$  with the condition  $n_a n^a = 0$  imposed by a Lagrange multiplier  $\lambda(x) g_{ab} n^a n^b$  added to  $\mathfrak{S}[n^a]$ . An elementary calculation and use of generalized Bianchi identity and the condition  $\nabla_a T_b^a = 0$  leads us to [3, 4, 134, 135] the following construct on background geometry:

$$\mathcal{G}_b^a = \mathcal{R}_b^a - \frac{1}{2} \delta_b^a L = \frac{1}{2} T_b^a + \Lambda \delta_b^a \quad (263)$$

where  $\Lambda$  is an integration constant. These are precisely the field equations for gravity in a theory with Lanczos-Lovelock Lagrangian  $L$  with an undetermined cosmological constant  $\Lambda$  arising as an integration constant.

The thermodynamical potential corresponding to the density  $\mathfrak{S}$  can be obtained by integrating the density  $\mathfrak{S}[n^a]$  over a region of space or a surface etc. based on the context. The matter part of the  $\mathfrak{S}$ , proportional to  $T_{ab} n^a n^b$ , will pick out the contribution  $(\rho + p)$  for an ideal fluid, which is the enthalpy density. Multiplied by  $\beta = 1/T$ , this reduces to the entropy density because of Gibbs-Duhem relation. When the multiplication by  $\beta$  can

be reinterpreted in terms of integration over  $(0, \beta)$  of the time coordinate (in the Euclidean version of the local Rindler frame), the  $\mathfrak{S}$  can be interpreted as entropy and the integral over space coordinates alone can be interpreted as rate of generation of entropy. (This was the interpretation provided in Refs. [3, 4, 134, 135] but the final result is independent of this interpretation as long as suitable boundary conditions can be imposed). One can also think of  $\mathfrak{S}[n^a]$  as an effective Lagrangian for the collective variables  $n^a$  describing the deformations of null surfaces.

The gravitational entropy density in Eq. (262) can be expressed in terms of the Killing ( $\mathcal{L}_q g_{ab} = \nabla_a q_b + \nabla_b q_a$ ) and isoentropic ( $J_{ab} = \nabla_a q_b - \nabla_b q_a$ ) parts of the deformations introduced in Ref. [5] as:

$$-4P_{cd}^{ab}\nabla_a q^c\nabla_b q^d = P^{bijd}(\mathcal{L}_q g_{ij})(\mathcal{L}_q g_{bd}) - (1/2)P^{abcd}J_{ab}J_{cd} \quad (264)$$

So the gravitational entropy density has two parts: one contributed by the square of the Noether potential (which vanish for isoentropic deformations with  $q_a = \nabla_a \phi$ ) and another which depends on the change in the metric under the deformation (which will vanish for the Killing deformations). When  $q_j$  is a pure gradient,  $J_{ab}$  will vanish and one can identify the first term in Eq. (264) with a structure like  $Tr(K^2) - (Tr K)^2$ . On the other hand, when  $q_a$  is a local Killing vector, the contribution from  $S_{ij}$  to the entropy density vanishes and the entropy density is just the square of the Noether potential  $J_{ab}$ . For a general null vector, both the terms contribute to the entropy density. Variation of entropy density with respect to either of the two contributions (after adding suitable Lagrange multiplier to ensure vanishing of the other term) will also lead to the gravitational field equations.

In this approach, there arise several new features which are worth mentioning.

First, the extremum value of the thermodynamic potential, when computed on-shell for a solution with static horizon, leads to the Wald entropy[134]. This is a non-trivial consistency check on the approach because it was not designed to reproduce this result. When the field equations hold, the total entropy of a region  $\mathcal{V}$  resides on its boundary  $\partial\mathcal{V}$  which is yet another illustration of the holographic nature of gravity.

Second, the entropy functional in Eq. (262) is invariant under the shift  $T_{ab} \rightarrow T_{ab} + \rho_0 g_{ab}$  which shifts the zero of the energy density. This symmetry allows any low energy cosmological constant, appearing as a parameter in the variational principle, to be removed thereby alleviating the cosmological constant problem to a great extent [135, 136]. As far as we know, *this is the only way in which gravity can be made immune to the zero point level of energy density*. It is again interesting that our approach leads to this result even though it is not designed for this purpose. This works because the cosmological constant, treated as an ideal fluid, has zero entropy because  $\rho + p = 0$  and thus cannot affect the dynamics in this perspective, in which gravity responds to the entropy density rather than energy density.

Third, the *algebraic* reason for the idea to work is the identity:

$$4P_{ab}^{cd}\nabla_c n^a\nabla_d n^b = 2\mathcal{R}_{ab}n^a n^b + \nabla_c[4P_{ab}^{cd}n^a\nabla_d n^b] \quad (265)$$

which shows that, except for a boundary term, we are extremising the integral of  $(2\mathcal{R}_{ab} - T_{ab})n^a n^b$  with respect  $n^a$  subject to the constraint  $n_a n^a = 0$ . The algebra is trivial but not the underlying physical concept.

Fourth, the gravitational entropy density — which is the term in the integrand  $\mathfrak{S}_{grav} \propto$



$(-P_{ab}^{cd}\nabla_c n^a \nabla_d n^b)$  in Eq. (262) — also obeys the relation:

$$\frac{\partial \mathfrak{S}_{\text{grav}}}{\partial(\nabla_c n^a)} = -8(-P_{ab}^{cd}\nabla_d n^b) = \frac{1}{4\pi}(\nabla_a n^c - \delta_a^c \nabla_i n^i) \quad (266)$$

where the second relation is for Einstein's theory. This term is analogous to  $t_a^c = K_a^c - \delta_a^c K$  (where  $K_{ab}$  is the extrinsic curvature) that arises in the (1+3) separation of Einstein's equations. (More precisely, the projection to 3-space leads to  $t_a^c$ .) This term has the interpretation as the momentum conjugate to the spatial metric in (1+3) decomposition and Eq. (266) shows that the entropy density leads to a similar structure. That is, the canonical momentum conjugate to metric in the conventional approach and the momentum conjugate to  $n^a$  in  $\mathfrak{S}_{\text{grav}}$  are similar.

One final question which must be asked is what justifies the starting expression for entropy functional as the “right” one. This has no answer in absence of a complete quantum gravity model. The next best thing one can do, and as we have done repeatedly in this review, is to compare the on-shell value of the expression near a Rindler horizon with Wald entropy. This can be easily done, and the entropy functional described here can be shown to reproduce correctly the known expression for Wald entropy on-shell. The proof of this is straightforward and can be found in [134].

## 6 Conclusions

The Lanczos-Lovelock models provide a simple and elegant generalization of Einstein's theory of gravity to higher dimensions. If it turns out that we actually live in a spacetime with more than 4-dimensions, then Lanczos-Lovelock models probably will prove to be of practical utility in the study of nature. But even if this is not the case, Lanczos-Lovelock models are of great interest in providing a natural backdrop to test different conceptual and mathematical aspects of Einstein's theory. This has been the point of view taken up and illustrated throughout the review. On several occasions we have pointed out how looking at a generalized result in Lanczos-Lovelock theory allows us to obtain a better insight into the phenomena in Einstein's theory. Such a study also allows us to delineate special features and numerical coincidences which occur in the  $D = 4$  Einstein gravity from results of much greater generality.

One of the surprising developments in recent years is the effortless generalizations of many concepts related to thermodynamics of horizons from Einstein gravity to these models. Based on this one could also argue that the entire emergent paradigm of gravity possesses a much broader basis than one would have first guessed from the study of Einstein's theory. Because of its importance, we have used these topics as the focal point for this review. We hope it helps to attract more attention from the community to this fascinating subject.

## Acknowledgements

The review draws upon the work done by the authors in collaboration with Sanved Kolekar, Bibhas Majhi, Ayan Mukhopadhyay, Aseem Paranjape, Sudipta Sarkar, L. Sriramkumar and Alexander Yale. One of the authors (TP) thanks Sunu Engineer for discussions. The work of TP is partially supported by J.C. Bose research grant.



## Appendices

### A Algebraic aspects of certain tensors in higher curvature theories of gravity

We here give a detailed analysis of certain tensors defined in Sec. 2.1 which appear naturally while varying a general Lagrangian which is a function of (and *only* of) metric and Riemann tensor and clarify many of the issues listed in Sec. 2.1. We will prove some useful identities which are non-trivial in the sense that they do not emerge from the *algebraic* symmetry properties of tensors such as  $P_{abcd}$  defined in Sec. 2.1; instead they arise due to the *diffeomorphism covariance* of the Lagrangian. This leads to the fact that, when scalar quantities are constructed from tensorial objects (like in the case of  $L$  built from  $R_{abcd}$  and  $g^{ij}$ ), the derivatives of the scalar quantity with respect to these tensors must obey some identities. While a few of these results can be proved by other methods in special cases, the procedure outlined here, based on reference [7], is very powerful and readily adaptable to more general cases.

While varying  $L = L[g_{ab}, R_{abcd}]$  to obtain the equations of motion, we come across two tensors in the following partial derivatives

$$P^{abcd} = \left( \frac{\partial L}{\partial R_{abcd}} \right)_{g_{ij}} ; \quad P^{ab} = \left( \frac{\partial L}{\partial g_{ab}} \right)_{R_{ijkl}} \quad (\text{A.1})$$

The subscripts specify the variables being held fixed, which, of course is, what gives meaning to the partial derivatives. The  $P^{abcd}$  is what we called the *entropy tensor* of the theory [41, 3]. The derivatives appearing in Eq. (A.1) are defined within the subspace of infinitesimal deformations which preserve the algebraic symmetries of the independent variable. Therefore,  $P^{ab}$  is symmetric while  $P^{abcd}$  inherits the algebraic symmetries of the curvature tensor:

$$P^{abcd} = -P^{bacd} = -P^{abdc}; \quad P^{abcd} = P^{cdab}; \quad P^{a[bcd]} = 0 \quad (\text{A.2})$$

One can construct several such tensors — which have the properties in Eq. (A.2) — from the curvature tensor and the metric. But these properties alone do *not* guarantee that such a tensor can be expressed as the derivative of some scalar with respect to the curvature tensor. The fact that  $P^{abcd}$  has the form in Eq. (A.1) leads to several *further* properties and inter-relationships which form the main thrust of this appendix. In particular, we will prove the following results:

- (1) Let us construct the tensor

$$\mathcal{R}^{ab} \equiv P^{aijk} R^b_{ijk} \quad (\text{A.3})$$

which can be thought of generalized Ricci tensor in these theories. We will prove that  $P^{ab} = -2\mathcal{R}^{ab}$ . That is, the derivatives of an *arbitrary* scalar function with respect to the metric and the curvature are *not* independent but are connected by the relation:

$$\left( \frac{\partial L}{\partial g_{ab}} \right)_{R_{ijkl}} = -2R^b_{ijk} \left( \frac{\partial L}{\partial R_{aijk}} \right)_{g_{ab}} \quad (\text{A.4})$$

It follows from Eq. (A.4) that the tensor  $\mathcal{R}^{ab}$  is symmetric. (This was noted earlier in a different context in [137]). This result does not follow merely from the algebraic symmetries of  $P^{abcd}$  in Eq. (A.2) above. The nature of the proof which we give below shows that it holds *only because*  $P^{abcd}$  is expressible as a derivative of a scalar with respect to the curvature tensor.

(2) From Eq. (A.4) we can obtain several other relations connecting the partial derivatives when different independent variables like  $(g^{ab}, R_{abcd})$  or  $(g^{ab}, R^a_{bcd})$  or  $(g^{ab}, R^{ab}_{cd})$  are used to describe the system. First, if we use the pair  $(g^{ab}, R_{abcd})$  as independent variables, we get:

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} = -g_{ai}g_{bm} \left( \frac{\partial L}{\partial g_{ab}} \right)_{R_{abcd}} = 2\mathcal{R}_{im} \quad (\text{A.5})$$

Instead, if we use the pair  $(g^{ab}, R^a_{bcd})$  as independent variables, one can also show

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R^a_{bcd}} = \mathcal{R}_{im} \quad (\text{A.6})$$

and, more interestingly, if we use the pair  $(g^{ab}, R^{ab}_{cd})$  as independent variables, we obtain

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R^{ab}_{cd}} = 0 \quad (\text{A.7})$$

(3) Using the fact that  $P^{bcid}$  is anti-symmetric in  $b$  and  $c$ , one can show

$$\nabla_b \nabla_c P^{bcid} = \frac{1}{2} (\mathcal{R}^{id} - \mathcal{R}^{di}) = 0 \quad (\text{A.8})$$

where the last equality follows from the symmetry of  $\mathcal{R}^{ab}$ . Thus we get another surprising result that, for any scalar  $L$  built from curvature tensor and the metric, we have the identity

$$\nabla_a \nabla_b \left( \frac{\partial L}{\partial R_{abcd}} \right)_{g_{ij}} = 0 \quad (\text{A.9})$$

We shall now sketch a proof of these results.

*Proof of the identities:*

As we mentioned in section (2.1), all the identities we want to prove follow from diffeomorphism invariance of the Lagrangian. So let us begin with an infinitesimal diffeomorphism  $x^a \rightarrow x^a + \xi^a(x)$  which changes  $L$ ,  $g_{ab}$  and  $R_{ijkl}$  by infinitesimal amounts. The essential trick is to express the Lie derivative of  $L$  in two different ways and equate the results. First, because  $L$  is a scalar which depends on  $x^i$  only through  $g_{ab}(x)$  and  $R_{ijkl}(x)$ , we have:

$$\mathcal{L}_\xi L = \xi^m \nabla_m L = \xi^m P^{ab} \nabla_m g_{ab} + \xi^m P^{abcd} \nabla_m R_{abcd} = P^{abcd} \xi^m \nabla_m R_{abcd} \quad (\text{A.10})$$

because  $\nabla_m g_{ab} = 0$ . On the other hand, if we think of the change  $\delta L$  in  $L$ , due to small changes in  $\delta g_{ab} \equiv \mathcal{L}_\xi g_{ab}$  and  $\delta R_{ijkl} \equiv \mathcal{L}_\xi R_{ijkl}$ , we also have:

$$\mathcal{L}_\xi L = P^{ab} \mathcal{L}_\xi g_{ab} + P^{ijkl} \mathcal{L}_\xi R_{ijkl} \quad (\text{A.11})$$

We will now expand the right hand side of Eq. (A.11) bringing it into a form similar to the right hand side of Eq. (A.10) and equate the two expressions. That will lead to Eq. (A.4). The first term on the RHS of Eq. (A.11) is easily dealt with; using  $\mathcal{L}_\xi g_{ab} = \nabla_a \xi_b + \nabla_b \xi_a$  and the symmetry of  $P^{ab}$  we obtain

$$P^{ab} \mathcal{L}_\xi g_{ab} = 2P^{ab} \nabla_a \xi_b \quad (\text{A.12})$$

The term involving Lie derivative of curvature tensor gives (on contracting with  $P^{ijkl}$  and using Eq. (A.2))

$$P^{ijkl} \mathcal{L}_\xi R_{ijkl} = P^{ijkl} \xi^m \nabla_m R_{ijkl} + 2P^{ijkl} [(\nabla_i \xi^m) R_{mjkl} + (\nabla_k \xi^m) R_{ijml}] \quad (\text{A.13})$$

With some further index manipulations, this can be written as

$$P^{ijkl} \mathcal{L}_\xi R_{ijkl} = P^{ijkl} \xi^m \nabla_m R_{ijkl} + 4(\nabla_i \xi_m) \mathcal{R}^{im} \quad (\text{A.14})$$

where we have used the definition  $\mathcal{R}^{im} = P^{ijkl} R^m_{\ jkl}$ . Using Eq. (A.14) and Eq. (A.12) in Eq. (A.11) we get

$$\begin{aligned} \mathcal{L}_\xi L &= P^{ijkl} \xi^m \nabla_m R_{ijkl} + 2(\nabla_i \xi_m) [P^{im} + 2\mathcal{R}^{im}] \\ &= \mathcal{L}_\xi L + 2(\nabla_i \xi_m) [P^{im} + 2\mathcal{R}^{im}] \end{aligned} \quad (\text{A.15})$$

Since  $\xi_m$  is arbitrary, it follows that the term within square brackets in the last expression has to vanish, leading to the condition in Eq. (A.4):

$$P^{im} = \left( \frac{\partial L}{\partial g_{im}} \right)_{R_{abcd}} = -2P^{ijkl} R^m_{\ jkl} = -2\mathcal{R}^{im} \quad (\text{A.16})$$

Since the left-hand-side is symmetric, it follows that  $\mathcal{R}^{im}$  must be symmetric. *This symmetry is certainly not obvious from the definition of  $\mathcal{R}^{im}$  itself.*

Equation (A.8) can be obtained immediately from the above result, as follows:

$$\begin{aligned} 2\nabla_b \nabla_c P^{bcid} &= [\nabla_b, \nabla_c] P^{bcid} \\ &= R^b_{\ kbc} P^{kcid} + R^c_{\ kbc} P^{bkid} + R^i_{\ kbc} P^{bckd} + R^d_{\ kbc} P^{bcik} \\ &= -\mathcal{R}^{di} + \mathcal{R}^{id} = 0 \end{aligned} \quad (\text{A.17})$$

where the anti-symmetry and pair-exchange symmetries of  $P^{bcid}$  have been used.

*Important Corollaries:*

One can derive a few corollaries from Eq. (A.16) which are useful in expressing the *field equations* obtained from our Lagrangian. We first note that, since  $\delta g_{ab} = -g_{ai} g_{bj} \delta g^{ij}$ , we have:

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} = - \left( \frac{\partial L}{\partial g_{im}} \right)_{R_{abcd}} = 2P_i^{\ jkl} R_{mjkl} = 2\mathcal{R}_{im} \quad (\text{A.18})$$

Let us next consider how our results change if we use the pair  $(g^{im}, R^a_{\ bcd})$  or the pair  $(g^{ab}, R^{ij}_{\ kl})$  as the independent variables. Obviously that  $(\partial L / \partial g^{im})$  with  $R_{abcd}$  held fixed is *not* the same

as  $(\partial L/\partial g^{im})$  with  $R^a_{bcd}$  held fixed. To find their relationship, we note that

$$\begin{aligned} dL &= \left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} dg^{im} + P^{abcd} dR_{abcd} = \left( \frac{\partial L}{\partial g^{im}} \right)_{R^a_{bcd}} dg^{im} + P_a{}^{bcd} d(g^{ak} R_{kbcd}) \\ &= P^{mbcd} dR_{mbcd} + dg^{im} \left\{ \left( \frac{\partial L}{\partial g^{im}} \right)_{R^a_{bcd}} + \mathcal{R}_{im} \right\} \end{aligned} \quad (\text{A.19})$$

from which it follows that

$$\begin{aligned} \left( \frac{\partial L}{\partial g^{im}} \right)_{R^a_{bcd}} &= \left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} - \mathcal{R}_{im} \\ &= 2\mathcal{R}_{im} - \mathcal{R}_{im} = \mathcal{R}_{im} = P_i{}^{bcd} R_{mbcd} \end{aligned} \quad (\text{A.20})$$

More interestingly, if we work with  $(g^{ab}, R^{ij}_{kl})$  as the independent variables we get in place of Eq. (A.19) the result:

$$dL = \left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} dg^{im} + P^{abcd} dR_{abcd} = P^{abcd} dR_{abcd} + dg^{im} \left\{ \left( \frac{\partial L}{\partial g^{im}} \right)_{R^{ab}_{cd}} + 2\mathcal{R}_{im} \right\} \quad (\text{A.21})$$

leading to

$$\left( \frac{\partial L}{\partial g^{im}} \right)_{R^{ab}_{cd}} = \left( \frac{\partial L}{\partial g^{im}} \right)_{R_{abcd}} - 2\mathcal{R}_{im} = 0 \quad (\text{A.22})$$

This result shows that scalars built from  $g^{ij}$ ,  $R^{ab}_{cd}$  are actually independent of the metric tensor when  $R^{ab}_{cd}$  is held fixed.

We conclude by proving another important corollary: *The tensor  $\nabla_m \nabla_n P^{amnb}$  is symmetric in indices  $a$  and  $b$ .* This result is important to establish directly the symmetry of  $E_{ab}$  in Eq. (10). We have already proved that  $\mathcal{R}_{ab}$  is symmetric, which means the last term  $\nabla_m \nabla_n P^{amnb}$  in Eq. (10) must also be symmetric in  $a$  and  $b$ . To show this explicitly, consider the anti-symmetric part of this tensor

$$\nabla_m \nabla_n (P^{amnb} - P^{bmna}) = \nabla_m \nabla_n (P^{amnb} + P^{anbm}) = -\nabla_m \nabla_n P^{abmn} \quad (\text{A.23})$$

To obtain the first equality we have used the pair exchange symmetry and the anti-symmetry of the first two indices  $P^{ijkl}$ ; in arriving at the second equality we have used the cyclic relation  $P^{i[jkl]} = 0$ . It follows that the last term in Eq. (10) is symmetric only if the right hand side of Eq. (A.23) vanishes, which is assured by Eq. (A.17) and the symmetry of  $P^{abcd}$  under pair exchange.

Therefore, we see that each of the terms in Eq. (10) are individually symmetric. The results of this section are based on arguments which would apply to Lagrangians more general than the ones we have considered (and hence more general than the Lanczos-Lovelock models), presumably giving more identities etc when, say, covariant derivatives of Riemann are also involved. In some sense, the main basis for deriving such identities lies in the very first equation Eq. (A.10) which actually encapsulates the scalar nature of the Lagrangian  $L$ .

A more detailed discussion of the results proved here, and some further implications of the same, can be found in [7].

## B Construction of the Gauss-Bonnet lagrangian

Let us consider building a Lagrangian  $L_2$  which is a linear combination of quadratic terms in curvature  $R^2$ ,  $R_{ab}R^{ab}$  and  $R_{abcd}R^{abcd}$ . The general form for  $L_2$  can be written as

$$L_2 = \left[ \alpha R^2 + \beta R_{ab}R^{ab} + \gamma R_{abcd}R^{abcd} \right] \quad (\text{B.1})$$

and its variation will have the form

$$\delta(\sqrt{-g} L_2) = \sqrt{-g} E_2^{ab} \delta g_{ab} + \sqrt{-g} \nabla_k \delta v^k \quad (\text{B.2})$$

A long but straightforward calculation gives

$$E_2^{ab} = \frac{1}{2} g^{ab} L_2 - \left[ 2\alpha R R^{ab} + (2\beta + 4\gamma) R_{kc} R^{kacb} - 4\gamma R^{ac} R_c^b + 2\gamma R_{kcd}^a R^{bkcd} \right] \quad (\text{B.3})$$

$$+ \left[ (2\alpha + \beta + 2\gamma) \nabla^a \nabla^b R - \left( 2\alpha + \frac{\beta}{2} \right) g^{ab} \square R - (\beta + 4\gamma) \square R^{ab} \right] \quad (\text{B.4})$$

and

$$\delta v^k = M^{ackb} \nabla_c \delta g_{ab} \quad (\text{B.5})$$

$$2M^{cbak} = \left( 4\alpha R g^{a[c} g^{b]k} + \beta (2g^{bk} R^{ac} - g^{kc} R^{ba} - g^{ba} R^{kc}) + 4\gamma R^{bcka} \right) + (a \leftrightarrow b) \quad (\text{B.6})$$

For  $E_{ab}$  not to contain higher than second derivatives of the metric, one must demand

$$2\alpha + \beta + 2\gamma = 0 \quad (\text{B.7})$$

$$4\alpha + \beta = 0 \quad (\text{B.8})$$

$$\beta + 4\gamma = 0 \quad (\text{B.9})$$

The solution is  $\alpha = c_2 = \gamma$ ,  $\beta = -4c_2$ , where  $c_2$  is some arbitrary constant. We therefore arrive at the following final form for the full second order Lagrangian

$$L = (16\pi G)^{-1} [R + (32\pi^2)^{-1} \alpha L_2] + \text{boundary terms} \quad (\text{B.10})$$

where  $[\alpha] = (\text{length})^2$  and we have set  $c_2 = (32\pi^2)^{-1}$  (this convention follows from the fact that Gauss-Bonnet term will then be the Euler character in  $4D$ ) with

$$L_2 \propto [R^{abcd} R_{abcd} - 4R^{ab} R_{ab} + R^2] \quad (\text{B.11})$$

The Einstein-Gauss-Bonnet (EGB) system, with lagrangian given by a linear combination of Einstein-Hilbert lagrangian and  $L_2$  above is the simplest truncation of Lovelock model which has been well studied in the literature. For example, some consequences of the quasi-linearity of Einstein-Gauss-Bonnet (EGB) equations on the Cauchy problem as well as on ‘‘junction conditions’’ are discussed in [138], while [139] focusses on discussion of the conserved charges in EGB.

## C Proof of identities in Eqs. (146) - (148)

To prove the required identities, we will use the following two relations concerning Lanczos-Lovelock actions of order  $m$  in  $D$  dimensions (see Eqs. (30))

$$\begin{aligned} E_{j(m)}^{i(D)} &= -\frac{1}{2} \frac{1}{16\pi} \frac{1}{2^m} \delta_{jc_1 d_1 \dots c_m d_m}^{ia_1 b_1 \dots a_m b_m} R_{a_1 b_1}^{c_1 d_1} \dots R_{a_m b_m}^{c_m d_m} \\ E_{i(m)}^{i(D)} &= -\frac{D-2m}{2} L_m^D \end{aligned} \quad (C.1)$$

where  $E_{j(m)}^{i(D)}$  is the equation of motion tensor for Lanczos-Lovelock action; for e.g., for  $m=1$ ,  $E_{j(1)}^{i(D)} = (16\pi)^{-1} G_j^i$  where  $G_j^i$  is the Einstein tensor. Note that for maximally symmetric spacetime, which is the case we are considering for the  $(D-2)$  sub-manifold corresponding to the  $(D-2)$  dimensional horizon, we have  $G_2^2 = G_3^3 = \dots = G_{D-1}^{D-1}$  due to isotropy. Hence we can write the second relation in Eq. (C.1) for equation of motion tensor for the  $(D-2)$  dimensional maximally symmetric subspace as

$$E_{B(m)}^{A(D-2)} = -\frac{\delta_B^A}{2} \left[ \frac{(D-2)-2m}{D-2} \right] L_m^{D-2} \quad (C.2)$$

Using the above identity it is easy to prove Eq. (146) as

$$\begin{aligned} (Q_{0A}^{0B})_m^{(D-1)} &= \frac{1}{16\pi} \frac{1}{2^m} \delta_{[0A\mu_1 \dots \mu_{2m-2}]}^{[0B\nu_1 \dots \nu_{2m-2}]} \hat{R}_{\nu_1 \nu_2}^{\mu_1 \mu_2} \dots \hat{R}_{\nu_{2m-3} \nu_{2m-2}}^{\mu_{2m-3} \mu_{2m-2}} \\ &= \frac{1}{16\pi} \frac{1}{2^m} \delta_{[AA_1 \dots A_{2m-2}]}^{[BB_1 \dots B_{2m-2}]} \hat{R}_{B_1 B_2}^{A_1 A_2} \dots \hat{R}_{B_{2m-3} B_{2m-2}}^{A_{2m-3} A_{2m-2}} \\ &= -E_{A(m-1)}^{B(D-2)} = \frac{1}{2} \left( \frac{D-2m}{D-2} \right) L_{m-1}^{D-2} \delta_B^A \end{aligned} \quad (C.3)$$

where to obtain the second equality we have used Eq. (117) and replaced  $\hat{R}_{CD}^{AB}$  with the intrinsic curvature  $\hat{\hat{R}}_{CD}^{AB}$  defined completely in terms of the induced metric  $\gamma_{AB}$  of the horizon. Such a replacement is valid since the corresponding extrinsic curvature vanishes due to our assumption of staticity. The last equality then follows from using Eq. (C.2). To prove the second identity of Eq. (147) note that

$$\begin{aligned} (Q_{AC}^{BD})_m^{(D-1)} &= \frac{1}{16\pi} \frac{1}{2^m} \delta_{[AC\mu_1 \dots \mu_{2m-2}]}^{[BD\nu_1 \dots \nu_{2m-2}]} \hat{R}_{\nu_1 \nu_2}^{\mu_1 \mu_2} \dots \hat{R}_{\nu_{2m-3} \nu_{2m-2}}^{\mu_{2m-3} \mu_{2m-2}} \\ &= \frac{1}{16\pi} \frac{1}{2^m} \delta_{[ACA_1 \dots A_{2m-2}]}^{[BDB_1 \dots B_{2m-2}]} \hat{R}_{B_1 B_2}^{A_1 A_2} \dots \hat{R}_{B_{2m-3} B_{2m-2}}^{A_{2m-3} A_{2m-2}} \\ &= (Q_{AC}^{BD})_m^{(D-2)} = M_m^{(D-2)} \delta_{AC}^{BD} \end{aligned} \quad (C.4)$$

where to obtain the second equality we have used  $\hat{\hat{R}}_{\rho\nu}^{\hat{\hat{\mu}}} = 0$  which is true due to (i) static nature of the background geometry and (ii) maximal symmetry of the  $(D-2)$  dimensional subspace and we have again replaced  $\hat{R}_{CD}^{AB}$  with the intrinsic curvature  $\hat{\hat{R}}_{CD}^{AB}$ . Further, in the last equality we have again exploited maximal symmetry to get a  $\delta_{AC}^{BD}$  proportionality with a

proportionality constant  $M_m^{(D-2)}$  which we can determine by contracting the above equation with  $\hat{\hat{R}}_{CD}^{AB}$  to get

$$L_m^{(D-2)} = (Q_{AC}^{BD})_m^{(D-2)} \hat{\hat{R}}_{BD}^{AC} = M_m^{(D-2)} \delta_{AC}^{BD} \hat{\hat{R}}_{BD}^{AC} = 2M_m^{(D-2)} D^{-2} R \quad (C.5)$$

This completes the proof of identity in Eq. (147). To prove the third identity of Eq. (148) we again use the same procedure and write

$$\begin{aligned} \delta_{[A0CE\mu_1\cdots\mu_{2m-4}]}^{[B0DF\nu_1\cdots\nu_{2m-4}]} \hat{R}_{\nu_1\nu_2}^{\mu_1\mu_2} \cdots \hat{R}_{\nu_{2m-5}\nu_{2m-4}}^{\mu_{2m-5}\mu_{2m-4}} &= \delta_{[ACEA_1\cdots A_{2m-4}]}^{[BDFB_1\cdots B_{2m-4}]} \hat{R}_{B_1B_2}^{A_1A_2} \cdots \hat{R}_{B_{2m-5}B_{2m-4}}^{A_{2m-5}A_{2m-4}} \\ &= N_{m-1}^{(D-2)} \delta_{ACE}^{BDF} \end{aligned} \quad (C.6)$$

with the proportionality constant  $N_{m-1}^{(D-2)}$  determined by contracting both sides of the above expression by  $\hat{\hat{R}}_{CD}^{AB}$  and then taking a trace to get

$$(16\pi)2^{m-1} \frac{D-2m}{D-2} L_{m-1}^{D-2} = N_{m-1}^{(D-2)} (16\pi)2^1 \frac{D-4}{D-2} L_1^{D-2} \quad (C.7)$$

This completes the proof of identity in Eq. (148).

## References

- [1] Lanczos C., *Z. Phys.* **73** 147 (1932); Lanczos C., *Annals Math.* **39** 842 (1938);
- [2] Lovelock D., *J. Math. Phys.* **12** 498 (1971).
- [3] Padmanabhan T., *Thermodynamical Aspects of Gravity: New insights*, Rep. Prog. Phys. **73** 046901 (2010) [arXiv:0911.5004].
- [4] T. Padmanabhan, *Lessons from Classical Gravity about the Quantum Structure of Space-time*, *J.Phys. Conf.Ser.* **306** 012001 (2011) [arXiv:1012.4476].
- [5] T. Padmanabhan, *Structural Aspects Of Gravitational Dynamics And The Emergent Perspective Of Gravity*, AIP Conf. Proc., **1483**, 212-238 (2012) [arXiv:1208.1375].
- [6] Padmanabhan T 2010 *Gravitation: Foundations and Frontiers*, (Cambridge University Press UK).
- [7] T. Padmanabhan, *Phys.Rev.*, **D 84** 124041 (2011) [arXiv:1109.3846].
- [8] Q. Exirifard, M. M. Sheikh-Jabbari, *Phys.Lett.* **B 661**, 158-161 (2008) [arXiv:0705.1879].
- [9] Monica Borunda, Bert Janssen, Mar Bastero-Gil, *JCAP*, 0811:008 (2008) [arXiv:0804.4440].
- [10] N. Dadhich, J. Pons, *Phys. Lett.* **B 705**, 139-142 (2011) [arXiv:1012.1692]
- [11] Naresh Dadhich, *Pramana*, **74**, 875-882 (2010) [arXiv:0802.3034]; D. Kastor, *Class. Quan. Grav.* **29**, pp. 155007 (2012) [arXiv:1202.5287].

- [12] N. Dadhich, S. Ghosh, S. Jhingan, *Phys. Lett.* **B 711**, 196-198 (2012) [arXiv:1202.4575]
- [13] S. Deser and R. Jackiw, *Int. J. Mod. Phys.* **B10** (1996) [hep-th/9510145].
- [14] N. Kiriushcheva and S. Kuzmin (2006) [hep-th/0602042].
- [15] N. Kiriushcheva and S. Kuzmin, *Mod. Phys. Lett.* **A21** (2006) [hep-th/0510260].
- [16] Alexandre Yale, T. Padmanabhan, *Gen.Rel.Grav.*, **43**, 1549-1570 (2011) [arXiv:1008.5154].
- [17] C. Cherubini, D. Bini, S. Capozziello and R. Ruffini, *Int. J. Mod. Phys.* **D11** (2002) [gr-qc/0302095]; see eq (11).
- [18] Efstratios Tsantilis, Roland A. Puntigam, Friedrich W. Hehl, *A quadratic curvature Lagrangian of Pawowski and Raczka: a finger exercise with MathTensor* [gr-qc/9601002]; see eq.(15).
- [19] T. Eguchi, P. Gilkey, A. Hanson, *Physics Reports*, **6**, 213-393 (1980).
- [20] A. Bidal, *Lectures on Anomalies* (2008) [arXiv:0802.0634].
- [21] Sanved Kolekar, T. Padmanabhan, *Phys.Rev.*, **D 82**,024036 (2010) [arXiv:1005.0619]
- [22] Gibbons, G. W., and Hawking, S. W. (1977), *Phys. Rev.*, **D 15**, 2738; York, J. W., Jr. (1988), *Boundary Terms in the Action Principles of General Relativity*, in *Between quantum and cosmos: Studies and essays in honor of John Archibald Wheeler*, eds. W. H., Zurek, A., van der Merwe, and W. A., Miller (Princeton Univ. Press), 246.
- [23] C. Teitelboim, J. Zanelli (1987), *Class. Quantum Grav.*, **4**, L125-L129.
- [24] S. Deser, J. Franklin, *Class. Quantum Grav.*, **29** 072001 (2012) [arXiv:1110.6085].
- [25] O. Miskovic, R. Olea, *JHEP* 0710:028, (2007) and references cited therein.
- [26] Misner, Thorne, Wheeler, *Gravitation*, W. H. Freeman and Company (1973).
- [27] Stephen C. Davis, *Phys.Rev.*, **D67**, 024030 (2003) [arXiv:hep-th/0208205]; Nathalie Deruelle and Cristiano Germani, *Nuovo Cim.*, **B118**, 977-988 (2003) [arXiv:gr-qc/0306116]; Claude Barrabes and W. Israel, *Phys.Rev.*, **D71**, 064008 (2005) [arXiv:gr-qc/0502108].
- [28] R. Myers, *Phys. Rev.*, **D 36**, 392 (1987).
- [29] A. Mukhopadhyay, T. Padmanabhan, *Phys.Rev.*, **D 74**, 124023 (2006) [hep-th/0608120].
- [30] T. Padmanabhan, *Phys. Reports* **406**, 49 (2005) [gr-qc/0311036].
- [31] D. G. Boulware, S. Deser, *Phys. Rev. Lett.*, **55**, 2656 (1985).



- [32] S. Deser and B. Tekin, *Class. Quant. Grav.* **20** 4877 (2003) [arXiv:gr-qc/0306114]; S. Deser and J. Franklin, *Class. Quant. Grav.* **22** L103 (2005) [arXiv:gr-qc/0506014]; R. Zegers, *J. Math. Phys.* **46** 072502 (2005) [arXiv:gr-qc/0505016]; C. Charmousis and J. Dufaux, *Class. Quant. Grav.* **19** 4671 (2002) [arXiv:hep-th/0202107]; H. Maeda and M. Nozawa, *Phys. Rev.* **D77** 064031 (2008) [arXiv:0709.1199].
- [33] Robert C. Myers, *Black holes, gravitational radiation and the universe*, Eds. Iyer, B.R. et al, 121-136 (Springer, 1998), [gr-qc/9811042].
- [34] Cecilia Garraffo, Gaston Giribet, *Mod.Phys.Lett.*, **A23**, 1801-1818 (2008) [arXiv:0805.3575].
- [35] R. Cai, K. Soh, *Phys. Rev.* , **D 59**, 044013 (1999) [gr-qc/9808067]; R. Cai, *Phys. Lett.* **B 582** 237-242 (2004) [hep-th/0311240]; R. Cai, N. Ohta, *Phys. Rev.* **D 74** 064001 (2006) [hep-th/0604088]; R. Cai, L. Cao, Y. Hu, S. P. Kim, *Phys. Rev.* **D 78** 124012 (2008) [arXiv:0810.2610];
- [36] N. Deruelle, L. Faria-Busto , *Phys. Rev.*, **D 41**, 3696 (1990).
- [37] J. D. Bekenstein, *Phys. Rev. D* **7**, 2333 (1973).
- [38] S. W. Hawking, *Nature* **248**, 30 (1974).
- [39] Davies P. C. W., *J. Phys. A* **8** 609–616 (1975); Unruh W. G., *Phys. Rev. D* **14** 870 (1976).
- [40] R. C. Myers, J. Z. Simon, *Phys. Rev.*, **D 38**, 2434 (1988).
- [41] Wald R. M., *Phys. Rev. D* **48** 3427 (1993) [gr-qc/9307038]; Iyer V., Wald R. M., *Phys. Rev. D* **52** 4430 (1995) [gr-qc/9503052].
- [42] M. Visser, *Phys. Rev. D* **48**, 583 (1993) [hep-th/9303029].
- [43] M. Visser, *Phys. Rev. D* **48**, 5697 (1993) [hep-th/9307194].
- [44] J. Lee, R. M. Wald, *J. Math. Phys.* **31**, 725 (1990).
- [45] S. N. Solodukhin, *Living Rev. Relativity* **14**, 8 (2011)
- [46] T. Padmanabhan, *Phys. Rev. Letts*, **78**, 1854 (1997) [hep-th/9608182]; T. Padmanabhan, *Phys. Rev.* **D 57** , 6206 (1998).
- [47] T. Padmanabhan, *Phys.Rev.*, **D 82**, 124025 (2010) [arXiv:1007.5066]
- [48] K. Srinivasan, T.Padmanabhan, *Phys. Rev.* **D 60** , 24007 (1999) [gr-qc/9812028]
- [49] S. Shankaranarayanan, K. Srinivasan and T. Padmanabhan, *Mod. Phys. Letts.*, **16**, 571 (2001) [gr-qc/0007022]
- [50] L. Vanzo, G. Acquaviva, R. Di Criscienzo, *Tunnelling Methods and Hawking’s radiation: achievements and prospects* [arXiv:1106.4153]

- [51] M. Parikh, F. Wilczek, *Phys. Rev. Letts.*, **85**, 5042 (2000) [hep-th/9907001]
- [52] S. Sarkar, D. Kothawala, *Phys. Lett. B* **659**, 683 (2008) [arXiv:0709.4448]
- [53] T. Padmanabhan, *Phys. Rept.* **188**, 285 (1990).
- [54] F. Pretorius, D. Vollick and W. Israel (1998), *Phys. Rev.*, **D 57**, 6311 [arXiv:9712085]
- [55] J. Oppenheim, *Phys. Rev.*, **D 65**, 024020 (2001) [arXiv:0105101]
- [56] S. Kolekar, T. Padmanabhan, *Phys. Rev.*, **D 83**, 064034 (2011) [arXiv:1012.5421]
- [57] R. Tolman and Paul Ehrenfest, *Phys. Rev.*, **36**, 1791-1798 (1930).
- [58] S. Kolekar and T. Padmanabhan, *Phys. Rev.*, **D 85**, 024004 (2012) [arXiv:1109.5353].
- [59] Thorne K S et al., *Black Holes: The Membrane Paradigm* (Yale University Press, 1986).
- [60] Sanved Kolekar, Dawood Kothawala, T. Padmanabhan, *Phys. Rev.*, **D 85**, 064031 (2012) [arXiv:1111.0973].
- [61] Maulik Parikh and Frank Wilczek, *Phys. Rev.*, **D 58**, 064011 (1998) [arXiv:gr-qc/9712077].
- [62] Ted Jacobson, Arif Mohd, Sudipta Sarkar, *The Membrane Paradigm for Gauss-Bonnet gravity* (2011) [arXiv:1107.1260].
- [63] Fu-Wen Shu, *Phys. Lett.*, **B685**, 325-328 (2010) [arXiv:0910.0607].
- [64] Xian Camanho and Jose Edelstein, *Causality in AdS/CFT and Lovelock theory* (2010) [arXiv:0912.1944].
- [65] Irene Bredberg, Cynthia Keeler, Vyacheslav Lysov and Andrew Strominger, *JHEP*, **1103**, 141 (2011) [arXiv:1006.1902].
- [66] P. Kovtun, D. T. Son and A. O. Starinets, *JHEP*, **0310**, 064 (2003) [arXiv:hep-th/0309213]; P. Kovtun, D. T. Son and A. O. Starinets, *Phys. Rev. Lett.*, **94**, 111601 (2005) [arXiv:hep-th/0405231].
- [67] Mauro Brigante, Hong Liu, Robert Myers, Stephen Shenker and Sho Yaida, *Phys. Rev.*, **D77**, 126006 (2008) [arXiv:0712.0805].
- [68] Xian Camanho, Jose Edelstein and Miguel Paulos, *JHEP*, **1105**, 127 (2011) [arXiv:1010.1682].
- [69] S. W. Hawking, *Commun. Math. Phys.* **25**, 152 (1972).
- [70] T. Jacobson, G. Kang, R. C. Myers, *Phys. Rev.* **D52**, 3518-3528 (1995) [gr-qc/9503020].
- [71] Sanved Kolekar T. Padmanabhan, , Sudipta Sarkar, *Phys. Rev.*, **D 86**, 021501(R) (2012) [arXiv:1201.2947].
- [72] A. Chatterjee and S. Sarkar, *Phys. Rev. Lett.* **108**, 091301 (2012) [arXiv:1111.3021].

- [73] R. M. Wald, “General Relativity,” *Chicago, Usa: Univ. Pr. ( 1984) p. 491*.
- [74] S. W. Hawking, G. F. R. Ellis, “The Large scale structure of space-time,” Cambridge University Press, Cambridge, 1973.
- [75] T. Jacobson and R. C. Myers, Phys. Rev. Lett. **70**, 3684 (1993) [arXiv:hep-th/9305016].
- [76] S. Gao and R. M. Wald, Phys. Rev. D **64**, 084020 (2001) [arXiv:gr-qc/0106071].
- [77] J. D. Brown and M. Henneaux, Commun. Math. Phys. **104**, 207 (1986).
- [78] A. Strominger, JHEP **9802**, 009 (1998) [arXiv:hep-th/9712251].
- [79] J. L. Cardy, Nucl. Phys. B **270** 186 (1986).  
H. W. J. Bloete, J. L. Cardy and M. P. Nightingale, Phys. Rev. Lett. **56**, 742 (1986);
- [80] S. Carlip, Class. Quant. Grav. **15**, 3609 (1998) [arXiv:hep-th/9806026].
- [81] S. Carlip, Phys. Rev. Lett. **82**, 2828 (1999) [arXiv:hep-th/9812013].
- [82] S. Carlip, Class. Quant. Grav. **16**, 3327 (1999) [arXiv:gr-qc/9906126].
- [83] T. Regge and C. Teitelboim, Annals Phys. **88**, 286 (1974).  
J. D. Brown and M. Henneaux, J. Math. Phys. **27**, 489 (1986).
- [84] J. Katz, J. Bicak and D. Lynden-Bell, Phys. Rev. D **55**, 5957 (1997) [arXiv:gr-qc/0504041].
- [85] R. M. Wald and A. Zoupas, Phys. Rev. D **61**, 084027 (2000) [arXiv:gr-qc/9911095].
- [86] B. Julia and S. Silva, Class. Quant. Grav. **15**, 2173 (1998) [arXiv:gr-qc/9804029].
- [87] B. Julia and S. Silva, Class. Quant. Grav. **17** 4733 (2000) [arXiv:gr-qc/0005127].
- [88] S. Silva, Class. Quant. Grav. **18**, 1577 (2001) [arXiv:hep-th/0010098].
- [89] S. N. Solodukhin, Phys. Lett. B **454**, 213 (1999) [arXiv:hep-th/9812056].
- [90] R. Brustein, Phys. Rev. Lett. **86**, 576 (2001) [arXiv:hep-th/0005266].  
G. A. S. Dias and J. P. S. Lemos, Phys. Rev. D **74**, 044024 (2006) [arXiv:hep-th/0602144].
- [91] F. L. Lin and Y. S. Wu, Phys. Lett. B **453**, 222 (1999) [arXiv:hep-th/9901147].  
V. O. Solovev, Phys. Rev. D **61**, 027502 (2000) [arXiv:hep-th/9905220].  
M. Natsuume, T. Okamura and M. Sato, Phys. Rev. D **61**, 104005 (2000) [arXiv:hep-th/9910105].  
D. J. Navarro, J. Navarro-Salas and P. Navarro, Nucl. Phys. B **580**, 311 (2000) [arXiv:hep-th/9911091].  
J. I. Jing and M. L. Yan, Phys. Rev. D **62**, 104013 (2000) [arXiv:gr-qc/0004061].
- [92] S. Das, A. Ghosh and P. Mitra, Phys. Rev. D **63**, 024023 (2001) [arXiv:hep-th/0005108].

- [93] J. I. Jing and M. L. Yan, Phys. Rev. D **63**, 024003 (2001) [arXiv:gr-qc/0005105].  
H. Terashima, Phys. Lett. B **499**, 229 (2001) [arXiv:hep-th/0011010].  
M. Hotta, K. Sasaki and T. Sasaki, Class. Quant. Grav. **18**, 1823 (2001) [arXiv:gr-qc/0011043].
- [94] H. Terashima, Phys. Rev. D **64**, 064016 (2001) [arXiv:hep-th/0102097].
- [95] J. I. Koga, Phys. Rev. D **64**, 124012 (2001) [arXiv:gr-qc/0107096].
- [96] M. I. Park, Nucl. Phys. B **634**, 339 (2002) [arXiv:hep-th/0111224].
- [97] S. Carlip, Phys. Rev. Lett. **88**, 241301 (2002) [arXiv:gr-qc/0203001].
- [98] M. Cvitan, S. Pallua and P. Prester, Phys. Lett. B **546**, 119 (2002) [arXiv:hep-th/0207265].  
A. Giacomini and N. Pinamonti, JHEP **0302**, 014 (2003) [arXiv:gr-qc/0301038].  
N. Pinamonti and L. Vanzo, Phys. Rev. D **69**, 084012 (2004) [arXiv:hep-th/0312065];  
G. Kang, J. I. Koga and M. I. Park, Phys. Rev. D **70**, 024005 (2004) [arXiv:hep-th/0402113].  
A. Giacomini, Phys. Rev. D **70**, 044005 (2004) [arXiv:hep-th/0403219].  
S. Carlip, Class. Quant. Grav. **22**, 1303 (2005) [arXiv:hep-th/0408123].  
S. Carlip, Class. Quant. Grav. **22**, R85 (2005) [arXiv:gr-qc/0503022].  
S. Carlip, Phys. Rev. Lett. **99**, 021301 (2007) [arXiv:gr-qc/0702107].  
S. Carlip, Gen. Rel. Grav. **39**, 1519 (2007) [Int. J. Mod. Phys. D **17**, 659 (2008)] [arXiv:0705.3024 [gr-qc]].  
L. M. Cao, Y. Matsuo, T. Tsukioka and C. M. Yoo, Phys. Lett. B **679**, 390 (2009) [arXiv:0906.2267 [hep-th]].  
L. Rodriguez and T. Yildirim, Class. Quant. Grav. **27**, 155003 (2010) [arXiv:1003.0026 [hep-th]].  
B. K. Button, L. Rodriguez and C. A. Whiting, arXiv:1009.1661 [hep-th].  
H. Chung, Phys. Rev. D **83**, 084017 (2011) [arXiv:1011.0623 [gr-qc]].
- [99] M. Cvitan, S. Pallua and P. Prester, Phys. Lett. B **555**, 248 (2003) [arXiv:hep-th/0212029].  
M. Cvitan, S. Pallua and P. Prester, Phys. Rev. D **70**, 084043 (2004) [arXiv:hep-th/0406186].  
M. Cvitan and S. Pallua, Phys. Rev. D **71**, 104032 (2005) [arXiv:hep-th/0412180].
- [100] M. I. Park and J. Ho, Phys. Rev. Lett. **83**, 5595 (1999) [arXiv:hep-th/9910158].  
S. Carlip, Phys. Rev. Lett. **83**, 5596 (1999) [arXiv:hep-th/9910247].
- [101] M. I. Park and J. H. Yee, Phys. Rev. D **61**, 088501 (2000) [arXiv:hep-th/9910213].
- [102] O. Dreyer, A. Ghosh and J. Wisniewski, Class. Quant. Grav. **18**, 1929 (2001) [arXiv:hep-th/0101117].
- [103] S. Silva, Class. Quant. Grav. **19**, 3947 (2002) [arXiv:hep-th/0204179].
- [104] B. R. Majhi, T. Padmanabhan, Phys. Rev. D **85**, 084040 (2012) [arXiv:1111.1809].

- [105] G. Barnich and F. Brandt, Nuclear Physics **B 633**, 3 (2002).
- [106] Bibhas Ranjan Majhi, T. Padmanabhan, *Phys.Rev.*, **D 86**, , 101501 (2012) [arXiv:1204.1422].
- [107] S. Zhang, B. Wang, [arXiv:1212.6896].
- [108] Bibhas Ranjan Majhi, D. Kothawala, T. Padmanabhan; under preparation.
- [109] T. Padmanabhan, *Research in Astron. Astrophys.*, **12**, 891916 (2012) [arXiv:1207.0505].
- [110] A. J. M. Medved, D. Martin, M. Visser, *Class. Quantum Grav.* **21**, 3111 (2004); *Phys. Rev.*, **D. 70**, 024009 (2004).
- [111] D. Kothawala and T. Padmanabhan, *Phys. Rev. D* **79**, 104020 (2009) [arXiv:0904.0215 [gr-qc]].
- [112] A. Paranjape, S. Sarkar and T. Padmanabhan, *Phys. Rev.* **D 74**, 104015, (2006) [hep-th/0607240].
- [113] T. Clunan, S. Ross, D. Smith, *Class. Quantum Grav.* **21**, 3447 (2004).
- [114] T. Padmanabhan, *Class. Quantum Grav.* **19**, 5387 (2002) [gr-qc/0204019]; *AIP Conf.Proc.*, **861**, 179, (2006) [astro-ph/0603114].
- [115] T. Padmanabhan, A Physical Interpretation of Gravitational Field Equations, AIP Conference Proceedings, 1241, 93-108 (2010) [arXiv:0911.1403].
- [116] T. Padmanabhan, *Gen.Rel.Grav*, **44**, 2681 (2012) [arXiv:1205.5683].
- [117] M. Parikh, S. Sarkar, arXiv:0903.1176; R. Brustein, M. Hadad, *Phys. Rev. Lett.* **103**, 101301 (2009).
- [118] T. Padmanabhan, *Entropy density of spacetime and thermodynamic interpretation of field equations of gravity in any diffeomorphism invariant theory*, [arXiv:0903.1254]; T. Padmanabhan, *Int.Jour.Mod.Phys.* **D 18** , 2189-2193 (2009).
- [119] T. Jacobson, *Thermodynamics of Spacetime: The Einstein Equation of State*, *Phys. Rev. Lett.* **75**, 1260 (1995).
- [120] C. Eling, R. Guedens, T. Jacobson, *Non-equilibrium Thermodynamics of Spacetime*, *Phys. Rev. Lett.* **96**, 121301 (2006).
- [121] R. Guedens, T. Jacobson, S. Sarkar, *Phys.Rev.*, **D 85**, 064017 (2012) [arXiv:1112.6215].
- [122] D. Kothawala, *Phys. Rev.*, **D 83**, 024026 (2011) [arXiv:1010.2207].
- [123] Padmanabhan T., *Class. Quan. Grav.*, **21**, 4485 (2004) [gr-qc/0308070].
- [124] Padmanabhan T., *Mod. Phys. Lett.*, **A 25**, 1129 (2010) [arXiv:0912.3165].
- [125] Padmanabhan T., *Phys. Rev.*, **D 81**, 124040 (2010) [arXiv:1003.5665].

- [126] J.D. Bekenstein, *Lett. Nuovo Cimento*, **11**, 467 (1974); J.D. Bekenstein [gr-qc/9710076].
- [127] Dawood Kothawala, T. Padmanabhan, Sudipta Sarkar, *Phys. Rev.*, **D78**, 104018 (2008) [arXiv:0807.1481].
- [128] T. Padmanabhan, *Mod.Phys.Letts. A* **17**, 1147 (2002). [hep-th/0205278]; *Gen.Rel.Grav.*, **34** 2029-2035 (2002) [gr-qc/0205090]; *Brazilian Jour.Phys. (Special Issue)* **35**, 362 (2005) [gr-qc/0412068]; F. Caravelli, L. Modesto, [arXiv:1001.4364].
- [129] Shabar Hod, *Phys. Rev. Lett.* **81**, 4293, (1998).
- [130] M. Maggiore, *Phys. Rev. Letts.*, **100**, 141301 (2008).
- [131] T. Padmanabhan, *Class. Quan. Grav.*, **21**, L1 (2004) [gr-qc/0310027]. T.Roy Choudhury, T. Padmanabhan, *Phys. Rev.*, **D 69** 064033 (2004) [gr-qc/0311064].
- [132] A. Medved, D. Martin, M. Visser, *Class. Quant. Grav.*, **21** 1393-1406 (2004) [gr-qc/0310009].
- [133] T. Padmanabhan, *Int.Jour.Mod.Phys. D***13**, 2293-2298 (2004) [gr-qc/0408051].
- [134] Padmanabhan T., *Gen. Rel. Grav.* **40** 2031-2036 (2008); Padmanabhan T., Paranjape A., *Phys.Rev. D* **75** 064004 (2007) [gr-qc/0701003].
- [135] Padmanabhan T., *Gen.Rel.Grav.* **40** 529-564 (2008) [arXiv:0705.2533]; Padmanabhan, T. *Phys. Rept.* **380**, 235-320 (2003) [hep-th/0212290].
- [136] Padmanabhan, T., *Adv. Sci. Lett.*, **2**, 174 (2009) [arXiv:0807.2356]; *Class.Quan.Grav.*, **22** L107-L110, (2005) [hep-th/0406060].
- [137] J. Koga and K. Maeda, *Phys. Rev. D* **58**, 064020 (1998).
- [138] N. Deruelle, J. Madore, “On the quasi-linearity of the Einstein- “Gauss-Bonnet” gravity field equations” [gr-qc/0305004].
- [139] N. Deruelle, J. Katz, S. Ogushi, *Class. Quant. Grav.* **21**, 1971 (2004) [gr-qc/0310098].